

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents
United States Patent and Trademark
Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 16 June 2000 (16.06.00)	
International application No. PCT/FI99/00870	Applicant's or agent's file reference 31531
International filing date (day/month/year) 20 October 1999 (20.10.99)	Priority date (day/month/year) 23 October 1998 (23.10.98)
Applicant YLIHONKO, Kristiina et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:
08 May 2000 (08.05.00)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was
☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer Manu Berrod
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

Applicant's or agent's
file reference

31531

International application No.

PCT / F 1 9 9 / 0 0 8 7 0

INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

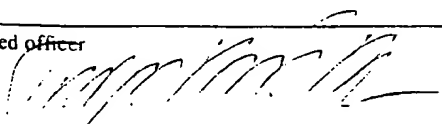
A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page <u>18</u> line <u>S 17-21</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ)	
Address of depositary institution (including postal code and country) Mascheroder Weg 1b, D-38124 Braunschweig, Germany	
Date of deposit 14 October 1998	Accession Number DSM 12451, DSM 12452
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect of those designations in which a European patent is sought, a sample of the deposited biological material will be made available as provided in Rule 28(3) EPC, until the publication of the mention of the grant of the European patent or for 20 years from the date of filing if the application is refused or withdrawn or deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only



This sheet was received with the international application

Authorized officer



For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

HAKEMUSNUMERO	LUOKITUS
982295	C 12 N 15/31, 15/52, 15/76, C 07 H 15/252, C 12 P 19/56, C 12 N 1/21 // (C 12 N 15/31, C 12 R 1:465) <input type="checkbox"/> jatkuu kääntöpuolella

TUTKITTU AINEISTO
Patenttivirastojen julkaisut FI, SE, NO, DK, DE, CH, EP, WO, GB, US:
FI, SE, NO, DK, julkaisut viraston julkaisukokoelmasta luokista C 12 N 15/31, 15/52, 15/76, C 07 H 15/252, C 12 P 19/56
Tietokannoista World Patent Index, EPO Documentation, Patent Abstracts of Japan, US full text (Epoque) <input type="checkbox"/> jatkuu kääntöpuolella
Muu aineisto
Tietokannoista Caplus, Biosis, Medline, Registry (STN) <input type="checkbox"/> jatkuu kääntöpuolella

VIITEJULKAISUT		
Kategoria*)	Julkaisun tunnistetiedot	Koskee vaatimuksia
A	Ylihonko, K. et al., WO 96/10581, C 07 K 14/36	1-15
A	Ylihonko, K. et al., Microbiology 142 (1996) 1965 - 1972	1-15
A	Ylihonko, K. et al., Mol. Gen. Genet. 251 (1996) 113 - 120	1-15
A	Torkkell, S. et al., Mol. Gen. Genet. 256 (1997) 203 - 209	1-15
<input type="checkbox"/> jatkuu kääntöpuolella		
*)	X Patentoitavuuden kannalta merkittävä julkaisu yksinään tarkasteltuna Y Patentoitavuuden kannalta merkittävä julkaisu, kun otetaan huomioon tämä ja yksi tai useampi samaan kategoriaan kuuluva julkaisu A Yleistä tekniikan tasoa edustava julkaisu, ei kuitenkaan patentoitavuuden este	
Päiväys	Tutkija	
5.7.1999	Stiina Kaikkonen	

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 31531	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/FI99/00870	International filing date (<i>day/month/year</i>) 20.10.1999	Priority date (<i>day/month/year</i>) 23.10.1998
International Patent Classification (IPC) or national classification and IPC ₇ C 07 K 14/36, C 12 N 15/31, C 12 P 19/56		
Applicant Galilaeus Oy et, al		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 3 sheets, including this cover sheet.

☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

 These annexes consist of a total of _____ sheets.

3. This report contains indications relating to the following items:
 - I ☒ Basis of the report
 - II ☐ Priority
 - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
 - IV ☐ Lack of unity of invention
 - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
 - VI ☐ Certain documents cited
 - VII ☐ Certain defects in the international application
 - VIII ☐ Certain observations on the international application

Date of submission of the demand 08.05.2000	Date of completion of this report 12.01.2001
Name and mailing address of the IPEA/SE Patent- och registreringsverket Box 5055 S-102 42 STOCKHOLM Facsimile No. 08-667 72 88	Authorized officer Patrick Andersson/EÖ Telephone No. 08-782 25 00

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/FI99/00870

I. Basis of the report

1. With regard to the **elements** of the international application:*

- ☒ the international application as originally filed
- ☐ the description:
pages _____, as originally filed
pages _____, filed with the demand
pages _____, filed with the letter of _____
- ☐ the claims:
pages _____, as originally filed
pages _____, as amended (together with any statement) under article 19
pages _____, filed with the demand
pages _____, filed with the letter of _____
- ☐ the drawings:
pages _____, as originally filed
pages _____, filed with the demand
pages _____, filed with the letter of _____
- ☐ the sequence listing part of the description:
pages _____, as originally filed
pages _____, filed with the demand
pages _____, filed with the letter of _____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language _____ which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☐ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. ☐ The amendments have resulted in the cancellation of:

- ☐ the description, pages _____
- ☐ the claims, Nos. _____
- ☐ the drawings, sheet/fig _____

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2 (c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

** Any replacement sheet containing such amendments must be referred to under item I and annexed to this report.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/FI99/00870

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**1. Statement**

Novelty (N)	Claims	<u>1-15</u>	YES
	Claims		NO
Inventive step (IS)	Claims	<u>1-15</u>	YES
	Claims		NO
Industrial applicability (IA)	Claims	<u>1-15</u>	YES
	Claims		NO

2. Citations and explanations (Rule 70.7)

The claimed invention relates to a gene cluster for the anthracycline biosynthetic pathway of *Streptomyces nogalater* comprising one 10kb fragment and one 7kb flanked BglII fragment of *S. nogalater* genome. The cluster, or parts thereof, can be used for synthesis of hybrid antibiotics.

The following document is considered relevant:

D1) Torkell S et al., "Characterization of *Streptomyces nogalater* genes encoding enzymes involved in the glycosylation steps in nogalamycin biosynthesis", Mol Gen Genet, 1997, vol 256, pages 203-209.

D1 discloses the characterization of *S. nogalater* genes involved in the glycosylation steps in nogalamycin biosynthesis. The genes are studied by inserting parts of a gene cluster into a strain of *S. galilaeus* (H039), thereby producing hybrid compounds. A sequence similarity search with the sequence for the presently claimed *snoal* gene indicates that *snoal* is known as a sub-sequence of sequence AF187532 related to D1.

In relation to the subject matter of the other claims D1 is considered to represent the general state of the art and has no particular relevance. Thus, the invention according to claims 1-15 is novel, inventive and industrially applicable.

P A T E N T C O O P E R A T I O N T R E A T Y

PCT

NOTIFICATION CONCERNING
SUBMISSION OR TRANSMITTAL
OF PRIORITY DOCUMENT

(PCT Administrative Instructions, Section 411)

From the INTERNATIONAL BUREAU

To:

OY JALO ANT-WUORINEN AB
Iso Roobertinkatu 4-6 A
FIN-00120 Helsinki
FINLANDE

Date of mailing (day/month/year) 25 November 1999 (25.11.99)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference 31531	
International application No. PCT/FI99/00870	International filing date (day/month/year) 20 October 1999 (20.10.99)
International publication date (day/month/year) Not yet published	Priority date (day/month/year) 23 October 1998 (23.10.98)
Applicant GALILAEUS OY et al	

1. The applicant is hereby notified of the date of receipt (except where the letters "NR" appear in the right-hand column) by the International Bureau of the priority document(s) relating to the earlier application(s) indicated below. Unless otherwise indicated by an asterisk appearing next to a date of receipt, or by the letters "NR", in the right-hand column, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).
2. This updates and replaces any previously issued notification concerning submission or transmittal of priority documents.
3. An asterisk(*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b). In such a case, **the attention of the applicant is directed** to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.
4. The letters "NR" appearing in the right-hand column denote a priority document which was not received by the International Bureau or which the applicant did not request the receiving Office to prepare and transmit to the International Bureau, as provided by Rule 17.1(a) or (b), respectively. In such a case, **the attention of the applicant is directed** to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

<u>Priority date</u>	<u>Priority application No.</u>	<u>Country or regional Office or PCT receiving Office</u>	<u>Date of receipt of priority document</u>
23 Octo 1998 (23.10.98)	982295	FI	16 Nove 1999 (16.11.99)

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No. (41-22) 740.14.35

Authorized officer

Catherine Massetti

Telephone No. (41-22) 338.83.38

PATENT COOPERATION TREATY

WO 00/24775
PCT/FI99/00870

12.05.2000

PCT

NOTICE INFORMING THE APPLICANT OF THE COMMUNICATION OF THE INTERNATIONAL APPLICATION TO THE DESIGNATED OFFICES

(PCT Rule 47.1(c), first sentence)

From the INTERNATIONAL BUREAU

To:

OY JALO ANT-WUORINEN AB
Iso Roobertinkatu 4-6 A
FIN-00120 Helsinki
FINLANDE

Date of mailing (day/month/year) 04 May 2000 (04.05.00)		
Applicant's or agent's file reference 31531		IMPORTANT NOTICE
International application No. PCT/FI99/00870	International filing date (day/month/year) 20 October 1999 (20.10.99)	
		Priority date (day/month/year) 23 October 1998 (23.10.98)
Applicant GALILAEUS OY et al		

1. Notice is hereby given that the International Bureau has communicated, as provided in Article 20, the international application to the following designated Offices on the date indicated above as the date of mailing of this Notice:

JP,US

In accordance with Rule 47.1(c), third sentence, those Offices will accept the present Notice as conclusive evidence that the communication of the international application has duly taken place on the date of mailing indicated above and no copy of the international application is required to be furnished by the applicant to the designated Office(s).

2. The following designated Offices have waived the requirement for such a communication at this time:

EP

The communication will be made to those Offices only upon their request. Furthermore, those Offices do not require the applicant to furnish a copy of the international application (Rule 49.1(a-bis)).

3. Enclosed with this Notice is a copy of the international application as published by the International Bureau on 04 May 2000 (04.05.00) under No. WO 00/24775

REMINDER REGARDING CHAPTER II (Article 31(2)(a) and Rule 54.2)

If the applicant wishes to postpone entry into the national phase until 30 months (or later in some Offices) from the priority date, a **demand for international preliminary examination** must be filed with the competent International Preliminary Examining Authority before the expiration of 19 months from the priority date.

It is the applicant's sole responsibility to monitor the 19-month time limit.

Note that only an applicant who is a national or resident of a PCT Contracting State which is bound by Chapter II has the right to file a demand for international preliminary examination.

REMINDER REGARDING ENTRY INTO THE NATIONAL PHASE (Article 22 or 39(1))

If the applicant wishes to proceed with the international application in the **national phase**, he must, within 20 months or 30 months, or later in some Offices, perform the acts referred to therein before each designated or elected Office.

For further important information on the time limits and acts to be performed for entering the national phase, see the Annex to Form PCT/IB/301 (Notification of Receipt of Record Copy) and Volume II of the PCT Applicant's Guide.

<p>The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland</p> <p>Facsimile No. (41-22) 740.14.35</p>	<p>Authorized officer</p> <p>J. Zahra</p> <p>Telephone No. (41-22) 338.83.38</p>
---	--

Continuation of Form PCT/IB/308

**NOTICE INFORMING THE APPLICANT OF THE COMMUNICATION OF
THE INTERNATIONAL APPLICATION TO THE DESIGNATED OFFICES**

Date of mailing (day/month/year) 04 May 2000 (04.05.00)	IMPORTANT NOTICE
Applicant's or agent's file reference 31531	International application No. PCT/FI99/00870
<p>The applicant is hereby notified that, at the time of establishment of this Notice, the time limit under Rule 46.1 for making amendments under Article 19 has not yet expired and the International Bureau had received neither such amendments nor a declaration that the applicant does not wish to make amendments.</p>	

PATENT COOPERATION TREATY

04.07.2000

From the INTERNATIONAL BUREAU

PCT

INFORMATION CONCERNING ELECTED
OFFICES NOTIFIED OF THEIR ELECTION

(PCT Rule 61.3)

To:

OY JALO ANT-WUORINEN AB
Iso Roobertinkatu 4-6 A
FIN-00120 Helsinki
FINLANDE

Date of mailing (day/month/year) 16 June 2000 (16.06.00)		
Applicant's or agent's file reference 31531		IMPORTANT INFORMATION
International application No. PCT/FI99/00870	International filing date (day/month/year) 20 October 1999 (20.10.99)	
Applicant GALILAEUS OY et al		Priority date (day/month/year) 23 October 1998 (23.10.98)

1. The applicant is hereby informed that the International Bureau has, according to Article 31(7), notified each of the following Offices of its election:

EP :AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE
National :JP,US

2. The following Offices have waived the requirement for the notification of their election; the notification will be sent to them by the International Bureau only upon their request:

None

3. The applicant is reminded that he must enter the "national phase" **before the expiration of 30 months from the priority date** before each of the Offices listed above. This must be done by paying the national fee(s) and furnishing, if prescribed, a translation of the international application (Article 39(1)(a)), as well as, where applicable, by furnishing a translation of any annexes of the international preliminary examination report (Article 36(3)(b) and Rule 74.1).

Some offices have fixed time limits expiring later than the above-mentioned time limit. For detailed information about the applicable time limits and the acts to be performed upon entry into the national phase before a particular Office, see Volume II of the PCT Applicant's Guide.

The entry into the European regional phase is postponed **until 31 months from the priority date** for all States designated for the purposes of obtaining a European patent.

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No. (41-22) 740.14.35	Authorized officer: Manu Berrod Telephone No. (41-22) 338.83.38
--	---



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C07K 14/36, C12N 15/31, C12P 19/56	A1	(11) International Publication Number: WO 00/24775 (43) International Publication Date: 4 May 2000 (04.05.00)
(21) International Application Number: PCT/FI99/00870 (22) International Filing Date: 20 October 1999 (20.10.99) (30) Priority Data: 982295 23 October 1998 (23.10.98) FI (71) Applicant (for all designated States except US): GALILAEUS OY [FI/FI]; Kairiskulmantie 10, FIN-20760 Piispanristi (FI). (72) Inventors; and (75) Inventors/Applicants (for US only): <u>YLIHONKO</u> , Kristiina [FI/FI]; Betonimiehenkatu 13, FIN-20780 Kaarina (FI). <u>TORKKELL</u> , Sirke [FI/FI]; Völppätie 3 B 39, FIN-20540 Turku (FI). <u>PALMU</u> , Kaisa [FI/FI]; Eerikinkatu 41 B 42, FIN-20100 Turku (FI). <u>HAKALA</u> , Juha [FI/FI]; Elinantie 2 A 9, FIN-20540 Turku (FI). (74) Agent: OY JALO ANT-WUORINEN AB; Iso Roobertinkatu 4-6 A, FIN-00120 Helsinki (FI).		(81) Designated States: JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> <i>With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.</i>
(54) Title: GENE CLUSTER INVOLVED IN NOGALAMYCIN BIOSYNTHESIS, AND ITS USE IN PRODUCTION OF HYBRID ANTIBIOTICS		
(57) Abstract The present invention relates to the gene cluster for nogalamycin biosynthesis derived from <i>Streptomyces nogalater</i> , and the use of the genes therein to obtain novel hybrid antibiotics for drug screening.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C07K 14/36, C12N 15/31, C12P 19/56	A1	(11) International Publication Number: WO 00/24775 (43) International Publication Date: 4 May 2000 (04.05.00)
(21) International Application Number: PCT/FI99/00870 (22) International Filing Date: 20 October 1999 (20.10.99) (30) Priority Data: 982295 23 October 1998 (23.10.98) FI (71) Applicant (for all designated States except US): GALILAEUS OY [FI/FI]; Kairiskulmantie 10, FIN-20760 Piispanristi (FI). (72) Inventors; and (75) Inventors/Applicants (for US only): YLIHONKO, Kristiina [FI/FI]; Betonimiehenkatu 13, FIN-20780 Kaarina (FI). TORKKELL, Sirke [FI/FI]; Völppätie 3 B 39, FIN-20540 Turku (FI). PALMU, Kaisa [FI/FI]; Eerikinkatu 41 B 42, FIN-20100 Turku (FI). HAKALA, Juha [FI/FI]; Elinantie 2 A 9, FIN-20540 Turku (FI). (74) Agent: OY JALO ANT-WUORINEN AB; Iso Roobertinkatu 4-6 A, FIN-00120 Helsinki (FI).		(81) Designated States: JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: GENE CLUSTER INVOLVED IN NOGALAMYCIN BIOSYNTHESIS, AND ITS USE IN PRODUCTION OF HYBRID ANTIBIOTICS (57) Abstract The present invention relates to the gene cluster for nogalamycin biosynthesis derived from <i>Streptomyces nogalater</i> , and the use of the genes therein to obtain novel hybrid antibiotics for drug screening.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Gene cluster involved in nogalamycin biosynthesis, and its use in production of hybrid antibiotics

Field of the invention

5

This invention relates to the gene cluster for nogalamycin biosynthesis derived from *Streptomyces nogalater*, and the use of the genes therein to obtain novel hybrid antibiotics for drug screening.

10 Background of the invention

Anthracyclines are antitumor antibiotics, mainly produced by *Streptomyces* sp. Daunomycin family of anthracyclines is commercially most important, since almost all of the around ten anthracyclines currently in clinical use, or in late clinical trials for cytotoxic
15 drugs, belong to this family. Despite the long history of anthracyclines, three decades or so, the studies on their biosynthesis are still going on, and there is further interest to obtain novel molecules for the development of cancer chemotherapeutics. A method currently used for finding novel molecules for drug screening is genetic engineering. Cloning the genes for anthracycline biosynthesis facilitates the production of hybrid
20 anthracyclines, as well as their use in combinatorial biosynthesis to generate novel molecules.

Nogalamycin, which was first described by Bhuyan and Dietz in 1965, is an anthracycline antibiotic produced by *Streptomyces nogalater*. It is highly active against tumor
25 cells, whereas toxic properties of this compound have prevented its progress to clinical trials (Bhuyan and Smith, 1975). However, menogaril (7-O-methylnogarol) is a semisynthetic derivative of nogalamycin, and its value in the treatment of cancer has been studied (e.g. Yoshida *et al.*, 1996), the interest being now mainly in Japan. Structurally nogalamycin (Fig. 1) differs from most other anthracyclines, as e.g. from
30 the daunomycin family, in two noteworthy features: (i) The stereochemistry at position nine is opposite, and (ii) it has a sugar moiety, in which nogalamine is attached at position 1 by a typical glycosidic bond, but it is also attached to carbon 2 by an

extraordinary C-C bond. Structural elucidation of nogalamycin was reported by Wiley *et al.* (1977). Furthermore, biosynthetic studies of nogalamycin have been published by Wiley *et al.* in 1978 giving information of the building blocks: The aglycone moiety is built from ten acetates; the neutral sugar, nogalose, is derived from glucose; and methyl groups of both of the sugars, nogalamine and nogalose, are transferred from methionine. The origin of nogalamine was not clearly solved by Wiley, but most probably nogalamine is also derived from glucose.

Molecular cloning of biosynthesis genes for anthracyclines has facilitated the studies on molecular genetics, providing tools for rational modifications of the structures, while also for surprising combinations with other antibiotics. Most of the interest has focused on daunomycin biosynthesis genes, as reported in several publications (Lomovskaya *et al.*, 1998; Rajgarhia and Strohl, 1997 and references therein). Some genes for aclacinomycin biosynthesis from *S. galilaeus* (Fujii and Ebizuka, 1997) and for rhodomycin biosynthesis from *S. purpurascens* (Niemi *et al.*, 1994) have been cloned as well. We have cloned the biosynthesis genes for nogalamycin, and successfully used the genes for producing hybrid anthracyclines. Most of the genes are involved in polyketide pathway, being responsible for the formation of a tricyclic intermediate, and they are reported in Ylihonko *et al.*, 1996a and b, and by Torkkell *et al.*, 1997. Despite the advances in molecular cloning, the biosynthetic pathway from glucose to sugars found in anthracyclines is still mainly hypothetical.

Regarding the genes for deoxyhexose pathway, Madduri *et al.* (1998) have reported that a gene derived from avermectin biosynthesis cluster caused the production of hybrid anthracyclines altering the sugar moiety when transferred into an *S. peucetius* mutant. The product obtained was epirubicin, a commercially important anthracycline. In this case a hydroxy group in the daunosamine moiety was in the opposite stereochemistry due to the action of an avermectin biosynthesis gene. *S. galilaeus* has been used as the host to prepare hybrid anthracyclines using the genes derived from rhodomycin pathway from *S. purpurascens* (Niemi *et al.*, 1994), and from nogalamycin biosynthesis cluster from *S. nogalater* (Ylihonko *et al.*, 1996a). The genes for nogalamycin pathway were used to generate the hybrid anthracycline production in *S. steffisburgensis* producing

typically steffimycin (Kunnari *et al.*, 1997). Previously, biosynthesis genes for actinorhodin have been expressed in *S. galilaeus* resulting in the formation of aloesaponarin (Strohl *et al.*, 1991). These hybrid compounds were modified in the aglycone moiety.

5 **Summary of the invention**

The present invention concerns a gene cluster of *Streptomyces nogalater*, most of the genes of the cluster being derived from the deoxyhexose pathway for nogalamine and nogalose. Expressing a DNA fragment of the said region in *S. galilaeus*, which produces
10 aclacinomycins, hybrid anthracyclines are obtained, wherein the aglycone moiety is derived from *S. galilaeus*, whereas the sugar moiety is characteristic neither to *S. nogalater* nor to *S. galilaeus*. Furthermore, when inserting the gene included in said cluster, encoding a cyclase for nogalamycin, into a suitable plasmid construction, nogalamycinone is obtained, which is the aglycone of nogalamycin. Since the stereo-
15 chemistry of nogalamycin differs from most other anthracyclines, using this gene enables the preparation of C-9 stereoisomers of the anthracycline molecules.

Detailed description of the invention

20 The experimental procedures of the present invention are methods conventional in the art. The techniques not described in detail here are given in the manuals by Hopwood *et al.* "Genetic manipulation of Streptomyces: a laboratory manual" The John Innes Foundation, Norwich (1985) and by Sambrook *et al.* (1989) "Molecular cloning: a laboratory manual". The publications, patents and patent applications cited herein are
25 given in the reference list in their entirety.

The present invention concerns particularly the gene cluster for nogalamycin biosynthesis (*Sno5*-cluster) causing the production of hybrid antibiotics with modifications in the sugar moiety. The invention concerns in specific the use of the genes for
30 nogalamine/nogalose biosynthesis to generate hybrid antibiotics modified in sugar moieties. The invention also concerns the use of a specific cyclase gene included in the

gene cluster of the invention, to generate the C-9 stereoisomers of typical anthracyclines.

5 The gene cluster according to the present invention is linked to the earlier reported clusters for nogalamycin biosynthesis. The starting point of the present invention was the gene cluster for nogalamycin chromophor (International Patent Application WO 96/10581). Subsequently, we have found some genes for the deoxyhexose pathway of nogalamycin biosynthesis (Torkkell *et al.*, 1997), and a part of the fragment comprising said genes was used to clone the genes for this invention.

10

The biosynthesis genes for nogalamycin can be isolated from *Streptomyces* sp., particularly from *S. nogalater*, which produces nogalamycin. Species which produce nogalamycin-like anthracyclines can also be used, e.g. *S. violaceochromogenes* producing arugomycin (Kawai *et al.*, 1987), or *S. avidinii* producing avidinorubicin (Aoki *et al.*,
15 1991).

Genomic DNA of a *Streptomyces* strain carrying the genes for nogalamycin biosynthesis is used in preparing a genomic library. Suitable gene fragments for cloning may be obtained by any frequently digesting restriction enzyme. Typically *Sau*3AI is used. The
20 isolated fragments could be inserted by ligation in any *Escherichia coli* vector such as a plasmid, a phagemid, a phage, or a cosmid. A cosmid vector is preferred since it enables the cloning of large DNA fragments. A cosmid vector such as pFD666 (ATCC No. 77286) is suitable for this purpose, as it enables cloning of the fragments of about 40 kb. The *Bam*HI site of pFD666, giving sticky ends to the *Sau*3AI fragments may be
25 used for cloning. Commercially available kits may be used to pack the DNA in phage particles. Various *E. coli* strains can be used for the infection by the DNA packed. An appropriate *E. coli* strain is, e.g. XL1Blue MRF', which is deficient in several restriction systems.

30 Using *E. coli* as a host strain for the genomic library, hybridization is an advantageous screening strategy. The probe for hybridization may be any known fragment derived from the nogalamycin gene cluster, but a short fragment of about 1 kb derived from one

end of the biosynthetic region previously cloned is preferred. Colonies for the genomic library are transferred for filter hybridization to membranes, preferably to nylon membranes. Since the average size for a genomic DNA fragment is 40 kb, 2300 colonies gave 99.99% probability to find the expanded region for nogalamycin biosynthesis. Any method for hybridization may be used but, in particular, the DIG System (Boehringer Mannheim, GmbH, Germany) is useful. Since the probe is homologous to the hybridized DNA, it is preferable to carry out the stringent washes of hybridization at 70°C in a low salt concentration according to Boehringer Mannheim's manual "DIG System User's Guide for Filter Hybridization". At least 80% homology is suggested to be needed for a DNA fragment to bind a probe in the conditions used for washes.

Using this protocol, seven clones out of about 5000 gave positive signals, and were picked up for DNA isolation. Restriction mapping is an appropriate technique for characterizing the clones. The positive clones may be digested with convenient restriction enzymes to demonstrate the physical linkage map of the DNA fragments. The cosmid used for cloning was a shuttle cosmid replicating in both *E. coli* and *Streptomyces* sp. However, the transfer of the recombinant cosmids in *S. lividans* TK24, which is a typically used laboratory strain in cloning *Streptomyces*, resulted in deletions, and was omitted. Instead, we rather used in the expression studies the plasmid pIJ486, a high copy number *Streptomyces* plasmid. However, any plasmid being able to stably replicate in *Streptomyces* may be used for this purpose.

Two *Bgl*II fragments of one of the clones were separately inserted into pIJ486 vectors, and the two plasmids obtained were transferred into a primary host, *S. lividans* TK24. The recombinant plasmids obtained (pSY42 and pSY43), containing a 10 kb and a 7kb fragment from *S. nogalater* genomic DNA, respectively, were isolated from the primary host and further introduced into other *Streptomyces* strains by protoplast transformation. The recombinant plasmid containing the 10 kb fragment caused the production of hybrid anthracyclines in the *S. galilaeus* mutant strain H039, which endogenously produces aklavinone-rhodinose-rhodinose-rhodinose. A few other *S. galilaeus* strains (H075, H026, H063) mutated in deoxyhexose pathway for sugars in aclacinomycin were used in

transformation, and new hybrid compounds were obtained. Since the structure of nogalamycin is almost unique among anthracyclines, the plasmids could be transferred to other anthracycline-producing strains, such as *S. peucetius*, which produces daunomycin, and *S. purpurascens*, which produces rhodomycins, to modify the structures of the characteristic antibiotics.

As the cloned cluster was linked to nogalamycin biosynthesis region already known, its ability to generate the modification in sugar moiety suggested the presence of the genes for deoxyhexose pathway. However, sequencing is necessary to deduce the function of the genes in the cluster cloned. The DNA fragments of 10 kb and 7 kb were further inserted into the plasmid pSL1190 for subcloning. Sequencing strategies such as a deletion set of the DNA fragments, shotgun cloning or primer walking could be used, but we prefer to use restriction fragments for subcloning. Using ABI PRISM system (Perkin-Elmer) for sequencing it is possible to get 500 to 700 bases per one reaction, which means that about 1 kb fragments sharing overlapping bases are needed for sequencing. For this purpose, 27 subclones were constructed.

Sequencing of the flanked *Bgl*III fragments consisting of about 16000 bp revealed 15 complete ORFs. The sequence analysis can be made by any computer based program, such as GCG (Madison, Wisconsin, USA) package. According to the present invention the putative gene functions as deduced from the sequence homology of those available in the libraries are

- aminotransferase (*snogI*), not completed
1. dTDP-glucose synthase (*snogJ*)
 - 25 2. aminomethyl transferase (*snogA*)
 3. polyketide cyclase, (*snoaM*)
 4. a gene of deoxyhexose pathway, unknown (*snogN*)
 5. hydroxylase, (*snoaG*)
 6. dTDP-4-dehydrorhamnose reductase (*snogC*)
 - 30 7. dTDP-glucose 4,6-dhydratase (*snogK*)
 8. NAME cyclase (*snoaL*)
 9. unknown (*snoK*)

10. glycosyl transferase, GTF (*snogD*)
11. unknown (*snoW*)
12. glycosyl transferase, GTF (*snogE*)
13. unknown (*snoL*)
- 5 14. unknown (*snoO*)
15. C-7 ketoreductase (*snoaF*)
- unknown (*snoN*), not completed

Gene designations: g means that the gene involved in biosynthesis of the glycosidic
 10 proportion including glycosyl transferases, whereas a points out that the gene is needed
 for the formation of the aglycone moiety.

Considering the proposed biosynthetic pathway for nogalamycin shown in Fig 3. we are
 able to cause several changes for the structures of antibiotics by the genes identified,
 15 including *snoaL*, responsible for the cyclization of the fourth ring of the aglycone
 moiety while determining the stereochemistry of the anthracyclinone, and the genes
 affecting the formation of nogalamine and nogalose (*snogJ*, *snogK*, *snogN*, *snogC*,
snogA), and, in addition, the genes responsible for joining the sugar residues to the
 aglycone moiety (*snogD* and *snogE*).

20

These genes could be separately inserted in a vector using suitable restriction sites, or
 by amplifying the genes by PCR. The fragments may contain an intrinsic promoter, or a
 promoter may be separately cloned. It is advantageous to use a vector carrying a
 promoter to allow expression of the genes in a *Streptomyces* strain. The plasmid
 25 pIJE486 contains a promoter *ermE* for erythromycin resistance gene, allowing constitut-
 ive expression of the genes inserted in a correct orientation. Special attention is drawn
 to the gene encoding a cyclase for the aliphatic ring, but any gene of said cluster may
 be expressed in *Streptomyces* hosts. The said cyclase converts the stereochemistry at C9
 of auramycinone in TK24, if inserted into the plasmid possessing the other genes for
 30 auramycinone biosynthesis, except the cyclase responsible for the typical
 stereochemistry of anthracyclines.

Streptomyces strains, in particular *S. galilaeus*, carrying the recombinant plasmids are cultivated in media wherein antibiotics are produced. The hybrid compounds are extracted with organic solvents from the culture broth, and the compounds are separated and purified using chromatographic techniques.

5

According to this invention *S. galilaeus* H039 carrying the plasmid pSY42 and designated as H039/pSY42 produces aklavinone-4'-epi-2-deoxyfucose in E1 medium supplemented with thiostrepton to give selection pressure for the plasmid containing strains.

10

S. lividans TK24 carrying the plasmid pSY15c containing the genes for the nogalamycin chromophore and the genes for a cyclase (*snoaL*) and a ketoreductase (*snoaF*), was cultivated in E1 medium supplemented with thiostrepton. The compound 9-epi-auramycinone was produced, and this structure is now called nogalamycinone. Any DNA fragment of the invention subcloned from a 17 kb nogalamycin biosynthesis region can be inserted in a vector replicating in *Streptomyces*, and the products may be produced by fermentation of the plasmid containing strains.

15

Brief description of the drawings

20

Fig. 1 shows the structures of nogalamycin, daunomycin and aclacinomycin.

Fig. 2 is a diagram of the gene cluster (*Sno5*) of the invention for nogalamycin biosynthesis.

25

Fig. 3 describes the proposed biosynthesis pathway for nogalamycin.

Fig. 4 shows a diagram of the plasmid pSY15c. The genes *snoaL* (**aL**) and *snoaF* (**aF**) shown black are inserted in the plasmid pSY15 to give pSY15c. **aL** represents a cyclase *snoaL* and **aF** is for C-7 ketoreductase *snoaF*. pSY15 (WO 96/10581) generates the production of a tricyclic intermediate for nogalamycin biosynthesis in *S. lividans*. The abbreviations **a1**, **a2** and **a3** refer to the

30

genes *snoa1*, *snoa2* and *snoa3*, respectively, for minimal PKS. *rA* is the *snorA* gene for an activator, *aB* is the *snoaB* gene for oxygenase, *aC* is the *snoaC* gene for methylase, *aD* is the *snoaD* gene for polyketide ketoreductase and *aE* is the *snoaE* for aromatase. *gF* (the *snogF* gene) and *gG* (the *snogG* gene) involved in the deoxyhexose pathway are not functional in the construct. *aph* is an aminoglycoside phosphotransferase gene, and *tsr* is a thiostreptone resistance gene.

Examples to further illustrate the invention are given hereafter.

10

EXPERIMENTAL

Materials used

Restriction enzymes used were purchased from Promega (Madison, Wisconsin, USA) or Boehringer Mannheim (Germany), and alkaline phosphatase from Boehringer Mannheim, and used according to the manufacturers' instructions. Proteinase K was purchased from Promega and lysozyme from Sigma (St. Louis, USA). HybondTM-N nylon membranes used in hybridization were purchased from Amersham (Buckinghamshire, England), DIG DNA Labelling Kit and DIG Luminescent Detection Kit from Boehringer Mannheim. Qiaquick Gel Extraction Kit from Qiagen (Hilden, Germany) was used for isolating DNA from agarose.

20

Bacterial strains and their use

- *Escherichia coli* XL1 Blue MRF' (Stratagene, La Jolla, CA) was used for cloning.
- 25 - *Streptomyces nogalater* ATCC 27451; the gene cluster of nogalamycin biosynthesis was cloned from this strain.

The host strains to express the genes cloned were:

- *Streptomyces lividans* TK24, also used as a primary host to clone DNA propagated in *E. coli*. The strain was provided by prof. Sir David Hopwood, John Innes Centre, UK.
- 30 - *Streptomyces galilaeus* H039, produces aklavinone-rhodinose-rhodinose-rhodinose
- *Streptomyces galilaeus* H026, produces aclacinomycin N, ACMN, (aklavinone-rhodosamine-2-deoxyfucose-rhodinose)

- *Streptomyces galilaeus* H063, produces aklavinone
- *Streptomyces galilaeus* H075, produces aklavinone-rhodosamine-2-deoxyfucose-2-deoxyfucose

5 The detailed description of the mutants H039 and H026 is given in Ylihonko *et al.* (1994) and of H075 in the FI patent application No. 981062 (Ylihonko *et al.*, 1998). H063 has not been described in the literature but it was obtained by NTG mutagenesis of *S. galilaeus*, and selected to be used as the host strain in the hybrid compound production, as it accumulates aklavinone without any sugar residues.

10

Plasmids

E. coli - *Streptomyces* shuttle cosmid pFD666 (ATCC 77286) was used for cloning the chromosomal DNA. *E. coli* cloning vectors pSL1190 (Pharmacia) and pUC19 were used for preparing the subclones.

15

pIJ486 is a high copy plasmid vector provided by prof. Sir David Hopwood, John Innes Centre, UK (Ward *et al.*, 1986)

20

pIJE486 is a vector containing *ermE* gene in the polylinker of pIJ486 (Bibb *et al.*, 1985).

pSY15 is a pIJ486 based plasmid construct, wherein the genes of polyketide pathway for nogalamycin biosynthesis were cloned (Ylihonko *et al.*, 1996a).

25 Nutrient media and solutions

For cultivation of *S. nogalater* for total DNA isolation TSB medium was used.

Lysozyme solution (0.3 M sucrose, 25 mM Tris, pH 8 and 25 mM EDTA pH 8) was used in isolation of total DNA. TE buffer (10 mM Tris, pH 8.0 and 1mM EDTA) was used to dissolve the DNA.

30

TRYPTONE-SOYA BROTH (TSB)

Per litre: Oxoid Tryptone Soya Broth powder 30 g.

ISP4

Bacto ISP-medium 4, Difco; 37 g/l.

E1 Per litre in tap water:

5	glucose	20 g
	soluble starch	20 g
	Farmamedia	5 g
	yeast extract	2.5 g
	K ₂ HPO ₄ •3H ₂ O	1.3 g
10	MgSO ₄ •7H ₂ O	1 g
	NaCl	3 g
	CaCO ₃	3 g

pH adjusted to 7.4 before autoclaving

15

General methods

NMR data was collected with a JEOL JNM-GX 400 spectrometer at the ambient temperature. ¹H and ¹³C NMR samples were internally referenced to TMS.

20

The anthracycline metabolites were detected by HPLC (LaChrom, Merck Hitachi, pump L-7100, detector L-7400 and integrator D-7500) using a LiChroCART RP-18 column (4.6x250mm). Acetonitrile:potassium hydrogen phosphate buffer (60 mM, pH 3.0 adjusted with citric acid) was used as the mobile phase. Gradient system starting from 25 65% to 30% of potassium dihydrogen phosphate buffer was used to separate the compounds. The flow rate was 1 ml/min and the detection was effected at 430 nm.

ISP4 plates supplemented with thiostrepton (50 µg/ml) were used to maintain the plasmid carrying cultures.

30

Example 1. Cloning the gene cluster for nogalamycin biosynthesis**1.1 Cosmid library**

For the isolation of total DNA, *Streptomyces nogalater* (ATCC 27451) was grown for 35 three days in 50 ml of TSB medium supplemented with 0.5% of glycine. The cells were harvested by centrifuging for 15 min at 3900 x g in 12 ml Falcon tubes, and the

cells were stored at -20°C . Cells from a 12 ml sample of the culture were used to isolate the DNA. 5 ml of lysozyme solution containing 5 mg of lysozyme/ml was added onto the cells, incubated for 20 min at 37°C . 500 μl of 10% SDS containing 0.7 mg of proteinase K was added onto the cells and incubated for 80 min at 62°C , another 500 μl of 10% SDS containing 0.7 mg of proteinase K was added, and incubation was continued for 60 min. The sample was chilled on ice and 600 μl of 3M NaAc, pH 5.8 were added, and the mixture was extracted with equilibrated phenol (Sigma). The phases were separated by centrifuging at $1400 \times g$ for 10 min. The DNA was precipitated from the water phase with equal volume of isopropanol to spool with a glass rod, and washed by dipping to 70% ethanol, air dried and dissolved in 500 μl of TE-buffer.

The chromosomal DNA was partially digested with *Sau3AI*. The DNA fragments were separated by agarose gel electrophoresis, and the fragments of 30 to 50 kb were cut from the 0.3% low gelling temperature SeaPlaque® agarose. The DNA bands were isolated from the gel by heating to 65°C , extracting with equal volume of equilibrated phenol, and the phases were separated by centrifuging for 15 min at $2500 \times g$. The phenol phase was extracted with TE buffer, centrifuged and the water phases were pooled. The DNA was precipitated by adding 0.1 volumes of NaAc, pH 5.8 and 2 volumes of ethanol at -20°C for 30 min, centrifuged for 30 min at 15 000 rpm in Sorvall RC5C centrifuge using SS-34 rotor with adapters for 10 ml tubes. The pellet was air dried and dissolved in 20 μl of TE buffer. The isolated fragments were ligated to pFD666 cosmid vector digested with *Bam*HI and dephosphorylated. The DNA was packed into phage particles, and infected to *E. coli* using Gigapack® III XL Packing Extract Kit according to the manufacturer's instructions.

25

1.2 Identification of the clones by hybridization

The infected cells were grown on LB plates containing 50 $\mu\text{g/ml}$ kanamycin and transferred to Hybond™-N nylon membranes (Amersham). The membranes were handled according to the protocol described in Boehringer Mannheim's manual "The DIG System User's Guide for Filter Hybridization". The probe used to screen the colonies for an expanded nogalamycin gene cluster was a 1.07 kb *Sac*I fragment from the cluster described earlier (Torkkell *et al.*, 1997). The plasmid carrying the probe was

30

digested with *SacI*, and the fragment was separated from the vector by agarose gel electrophoresis and isolated from the gel using Qiaquick Gel Extraction Kit (Qiagen). The probe was labelled by digoxigenin using random prime labelling system according to Boehringer Mannheim's manual "The DIG System User's Guide for Filter Hybrid-
5 ization". 5000 colonies were screened by hybridization at 70°C using the probe described. Positive colonies were detected using DIG Luminescent Detection Kit (Boehringer Mannheim). Seven colonies gave a positive signal. Cosmids from the positive clones were isolated from a 5ml culture by alkaline lysis method. Restriction
10 analysis showed that the cloned fragments overlapped each other representing at least 60 kb of the continuous DNA. The positive clones obtained were designated as pFDSno1 to pFDSno7.

1.3. Subcloning the fragments for sequencing

Clone No. 5, designated as pFDSno5, was digested with *BglII*, and for subcloning two
15 fragments of about 10 kb and 7 kb were isolated and ligated to pSL1190 digested with *BglII* and dephosphorylated. The plasmids obtained were named as pSn42 and pSn43, respectively. These two fragments cover the DNA region flanked to the previously characterized area of nogalamycin biosynthesis cluster. To determine the nucleotide
20 sequence of the whole 17 kb region cloned in pSn42 and pSn43 the convenient restriction sites were used to subclone the fragments to the vector pUC19 or pSL1190 giving 16 subclones from the insert of pSn42 and 11 subclones of pSn43.

E. coli XL1 Blue MRF' cells were cultivated overnight at 37 °C in 5 ml of LB-
medium supplemented with 50 µg/ml of ampicillin. To isolate plasmids for sequencing
25 reactions Wizard Plus Minipreps DNA Purification System kit of Promega, or Biometra silica spin plasmid miniprep kit of Biomedizinische Analytik GmbH were used according to the manufacturers' instructions.

DNA sequencing was performed using the automatic ABI DNA sequenator (Perkin-
30 Elmer) according to the manufacturer's instructions.

1.4 Sequence analysis and the deduced functions of the genes

Sequence analyses were effected using the GCG sequence analysis software package (Version 8; Genetics Computer Group, Madison, Wisconsin, USA). The translation table was modified to accept also GTG as a start codon. Codon usage was analysed using published data (Wright and Bibb 1992).

According to the CODONPREFERENCE program the sequenced DNA fragment (SEQ ID NO:1) contained 15 complete open reading frames (ORFs), and the 5' end of other two ORFs in the both ends of the fragment according to the invention. The functions of the genes were concluded by comparing the amino acid sequences translated from their base sequences to the known protein sequences in the data banks. The results are shown in Table 1. The positions given refer to the appended sequence listing. The amino acid sequences of the peptides are given in SEQ ID NO:2 to SEQ ID NO:18.

Table 1

Gene	Position	Amino acids (SEQ ID NO)	Deduced function	Remarks
<i>snogI</i>	-1027 compl	>342 (2)	aminotransferase	5' end
<i>snogJ</i>	1192-2073	293 (3)	dTDP-glucose synthase	
<i>snogA</i>	2106-2822 compl	238 (4)	aminomethyl transferase	
<i>snoaM</i>	2826-3800 compl	324 (5)	a polyketide cyclase	
<i>snogN</i>	3799-5025	408 (6)	<i>dnrQ</i> homology (Otten <i>et al.</i> , 1995), unknown	
<i>snoaG</i>	5088-6356	422 (7)	hydroxylase	
<i>snogC</i>	6334-7209 compl	291 (8)	dTDP-4-dehydrorhamnose reductase	
<i>snogK</i>	7245-8297 compl	350 (9)	dTDP-glucose-4,6-de-hydratase	
<i>snoaL</i>	8537-8941	134 (10)	NAME cyclase (nogalonic acid methyl ester)	
<i>snoK</i>	8992-9699	235 (11)	unknown	
<i>snogD</i>	9745-10917 compl	390 (12)	glycosyl transferase	
<i>snoW</i>	11057-11884	275 (13)	unknown	
<i>snogE</i>	11928-*	>424 (14)	glycosyl transferase	
<i>snoL</i>	13335-13754 compl	139 (15)	unknown	
<i>snoO</i>	13974-14441	155 (16)	homologous to <i>mtmX</i> of mithramycin cluster	
<i>snoaF</i>	14532-15377	281 (17)	C-7 ketoreductase analogous to aklaviketone ketoreductase	
<i>snoN</i>	15450-	>190 (18)	unknown	5' end

*: nucleotide sequence of about 100 bp, not known

1.5 Expression cloning

The 10 kb *Bgl*II fragment from pFDSno5 was cloned into the plasmid pIJ486 and the plasmid obtained was named as pSY42. Correspondingly, the 7 kb *Bgl*II fragment from pFDSno5 was cloned into the plasmid pIJE486, and the plasmid pSY43 was obtained.

- 5 Plasmid pSY42 was introduced into *S. lividans* strain TK24 by protoplast transformation, isolated from it and further introduced into *S. galilaeus* mutant H039, and after propagation in H039, transferred to other *S. galilaeus* mutants blocked in the deoxyhexose pathway for characteristic sugars of aclacinomycins (H075, H026, and H063). E1 medium was used for anthracycline production, and the products were extracted
- 10 from the culture with toluene:methanol (1:1) at pH 7. Anthracycline metabolites were analyzed by HPLC. The products of the mutants H039, H026, H063 and H075 carrying pSY42 differed from those obtained by the mutants without the plasmid.

- According to the sequence analysis pSY42 contained a cyclase designated as NAMEC
- 15 (nogalonic acid methyl ester cyclase), and in pSY43 a ketoreductase gene was identified. Expression constructions were prepared which contained all the genes needed for the formation of nogalamycin aglycone. A 1.4 kb *Bam*HI-*Sac*I fragment from pSY42 (containing NAMEC) and a 1.1 kb *Mlu*I-*Kpn*I fragment from pSY43 carrying the gene for a ketoreductase of C-7 keto group were ligated to pSY15 linearized by *Sac*I, to
- 20 form the plasmid pSY15c (Fig. 4). Plasmid pSY15c was introduced into *S. lividans* TK24, and the strain TK24/pSY15c was cultivated in E1 medium supplemented with thiostrepton. An aglycone compound was produced, and this structure is now called nogalamycinone.

25 Example 2. Compounds generated by the *sno5*-cluster

2.1 Production and purification of the products derived from H039/pSY42 and TK24/pSY15c

- The seed culture, 180 ml of E1 culture of the plasmid containing strain, H039/pSY42 or
- 30 TK24/pSY15c, was obtained by cultivating the strain in three 250 ml Erlenmeyer flasks containing 60 ml of E1 medium supplemented with thiostrepton (5 μ g/ml) for four days at 30°C, 330 rpm. The combined culture broths (180 ml) were used to inoculate 13 l of

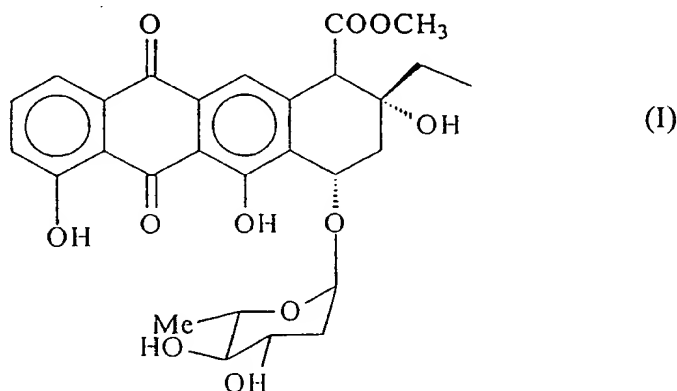
E1 medium in a fermentor (Biostat E). Fermentation was carried out for seven days at 28°C (330 rpm, aeration: 450 l/min).

The cells were harvested by centrifuging. 2.6 l of methanol was used to break the bacterial cells and to extract anthracycline metabolites accumulated. The anthracycline metabolites were extracted using 2 l of dichloromethane at pH 6. The organic layer was evaporated to dryness. The viscous residue was flashed through a polyamide (11) column using water:methanol from 1:9 to 0:10 as the eluent. Pooled fractions containing the compounds were further purified on a Merck–Hitachi HPLC using preparative reversed phase column (LichroCART RP–18, 5 μ m) with mobile phase acetonitrile:1 % AcOH in water (1:1). Evaporation of acetonitrile gave pure products as yellow powders dried under vacuum.

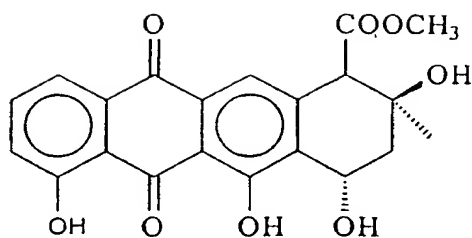
2.2 Structural elucidation of the compounds derived from H039/pSY42 and from TK24/pSY15c

NMR analysis included NON, BMC, NOE, DEPT and HMBC techniques. Protons were assigned using NOESY and 2D pTOCSY techniques and carbons using DEPT and HMBC techniques.

As deduced from the data given in Tables 2 and 3, the structures revealed were aklavinone–4'–epi–2–deoxyfucose from the culture of H039/pSY42, and 9–epi–auramycinone (=nogalamycinone) from the culture of TK24/pSY15c. The chemical structures of the compounds are shown below in Formula I and Formula II, respectively.



5



(II)

Deposited microorganisms

10

The following microorganisms were deposited according to the Budapest Treaty at Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D-38124 Braunschweig, Germany.

15	Microorganism	Accession number	Date of deposit
	<i>S. lividans</i> TK24/pSY42 carrying the plasmid pSY42	DSM 12451	14 October 1998
20	<i>S. lividans</i> TK24/pSY43 carrying the plasmid pSY43	DSM 12452	14 October 1998

Table 2. ^1H and ^{13}C assignments of the compound aklavinone-4'-*epi*-2-deoxyfucose (Formula I).

Site	^1H	^{13}C
1	7.74, 1H, dd, 7.5, 1.3	120.1
2	7.68, 1H, dd, 8.4, 7.5	137.3
3	7.27, 1H, dd, 8.3, 1.3	124.6
4	-	161.9
4-OH	11.70, 1H, s	-
4a	-	115.4
5	-	192.3
5a	-	114.4
6	-	162.4
6-OH	12.46, 1H, s	-
6a	-	130.9
7	5.18, 1H, dd, 4.3, 3.1	71.3
8A	2.51, 1H, dd, 15.0, 4.3	33.9
8B	2.32, 1H, dd, 15.0, 3.1	-
9	-	72.1
9-OH	4.58, 1H, s	-
10	4.02, 1H, s	56.9
10a	-	142.4
11	7.40, 1H, s	120.8
11a	-	133.1
12	-	180.7
12a	-	132.6
13A	1.73, 1H, dq, 14.2, 7.4	32.0
13B	1.51, 1H, dq, 14.2, 7.4	-
14	1.10, 3H, t, 7.4	6.7
15	-	171.1
16	3.69, 3H, s	52.5
1'	5.41, 1H, d, 3.5	101.7
2'a	1.75, 1H, ddd, 12.8, 11.2, 3.4	37.7
2'c	2.19, 1H, dd, 12.8, 5.3	-
3'	3.71, 1H, ddd, 12.0, 9.0, 5.3	69.0
4'	3.14, 1H, dd, 9.1, 9.0	78.1
5'	3.88, 1H, dq, 9.1, 6.2	68.8
6'	1.36, 3H, d, 6.2	17.6

Table 3. ^1H and ^{13}C assignments of 9-*epi*-auramycinone (Formula II).

Site	^1H	^{13}C
1	7.76, 1H, dd, 7.5, 1.2	119.8
2	7.67, 1H, dd, 8.3, 7, 5	137.4
3	7.28, 1H, dd, 8.3, 1.2	124.8
4	—	162.5
4-OH	11.86, 1H, s	—
4a	—	115.6
5	—	192.7
5a	—	114.6
6	—	160.9
6-OH	12.76, 1H, s	—
6a	—	134.1
7	5.40, 1H, t, 7.0	64.0
8A	2.66, 1H, dd, 13.9, 7.0	40.9
8B	1.89, 1H, dd, 13.9, 7.1	—
9	—	70.5
9-OH	3.49, 1H, brs	—
10	3.93, 1H, d, 0.8	56.0
10a	—	142.1
11	7.51, 1H, d, 0.8	120.1
11a	—	133.3
12	—	180.9
12a	—	132.1
13	1.44, 3H, s	28.7
14	—	173.0
15	3.90, 3H, s	52.6

References

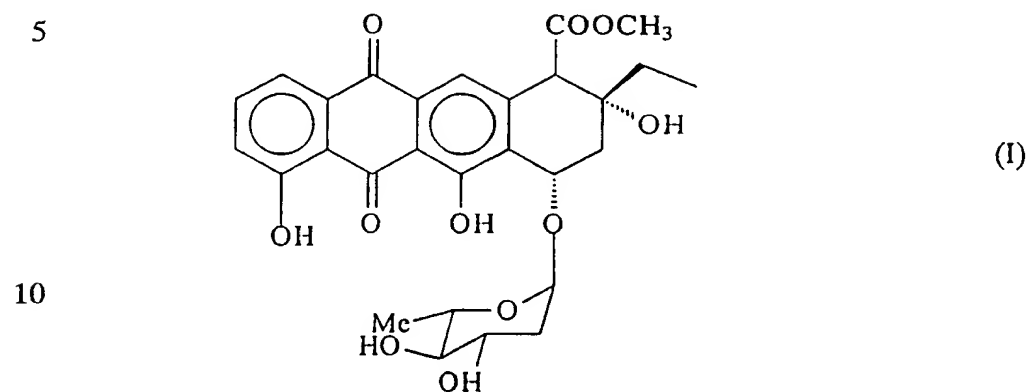
- Aoki, M., Shirai, H., Nakayama, N., Itezono, Y., Mori, Mitsuko, Satoh, T., Ohishima, S., Watanabe, J., and Yokose, K. 1991. Structural studies on avidinorubicin, a novel anthracycline with platelet aggregation inhibitory activity. *J. Antibiot.* **44**: 635–645.
- Bibb, M.J., Janssen, G.R., and Ward, J.M. 1985. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ErmE*) of *Streptomyces erythraeus*. *Gene* **38**: 215–226.
- Bhuyan, B.K., and Dietz, A. 1965. Fermentation, taxonomic and biological studies of nogalamycin. *Antimicrob. Ag. Chemother.* 1965: 836–844.
- Bhuyan, B.K., and Smith C.G. 1975. In: Antineoplastic and immunosuppressive agents, part II, pp. 623–632. Sartorell, A.C., and Johns, D.G. (eds), Springer-Verlag, Berlin.
- Fujii, I., and Ebizuka, Y. 1997. Anthracycline biosynthesis in *Streptomyces galilaeus*. *Chem. Rev.* **97**: 7: 2511–2523.
- Kawai, H., Hayakawa, Y., Nakagawa, M., Furihata, K., Shimazu, A., Seto, H. and Otake, N. 1987. Arugomycin, a new anthracycline antibiotic I. Taxonomy, fermentation, isolation and physico-chemical properties. *J. Antibiot.* **19**: 1266–1271.
- Kunnari, T., Tuikkanen, J., Hautala, A., Hakala, J., Ylihonko, K., and Mäntsälä, P. 1997. Isolation and characterization of 8-demethoxy steffimycins and generation of 2,8-demethoxy steffimycins in *Streptomyces steffisburgensis* by the nogalamycin biosynthesis genes. *J. Antibiot.* **50**: 496–501.
- Lomovskaya, N., Doi-Katayama, Y., Filippini, S., Nastro, C., Fonstein, L., Gallo, M., Colombo, A.L. and Hutchinson, C.R. 1998. The *Streptomyces peucetius dpsY* and *dnrX* genes govern early and late steps of daunorubicin and doxorubicin biosynthesis. *J. Bacteriol.* **180**: 9: 2379–2386.
- Madduri, K., Kennedy, J., Rivola, G., Inventi-Solari, A., Filippini, S., Zanuso, G., Colombo, A.L., Gewain, K.M., Occi, J.L., MacNeil, D.J., and Hutchinson, C.R. 1998. Production of antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of *S. peucetius*. *Nature Biotech.* **16**: 69–74.
- Niemi, J., Ylihonko, K., Hakala, J., Kopio, A., Pärssinen, R., and Mäntsälä, P. 1994. Hybrid anthracycline antibiotics: production of new anthracyclines by cloned genes from *Streptomyces purpurascens* in *Streptomyces galilaeus*. *Microbiol.* **140**: 1351–1358.
- Otten, S.L., Liu, X., Ferguson, J., and Hutchinson C.R. 1995. Cloning and characterization of the *Streptomyces peucetius dnrQS* genes encoding a daunosamine biosynthesis enzyme and a glycosyl transferase involved in daunorubicin biosynthesis. *J. Bacteriol.* **177**: 6688–6692.

- Rajgarhia, V.B., and Strohl, W.R. 1997. Minimal *Streptomyces* sp. strain C5 dauno-rubicin polyketide biosynthesis genes required for aklanonic acid biosynthesis. *J. Bacteriol.* **179**: 8: 2690–2696.
- Strohl, W.R., Bartel, P.L., Li, Y., Connors, N.C., and Woodman, R.H. 1991. Expression of polyketide biosynthesis and regulatory genes in heterologous streptomycetes. *J. Ind. Microbiol.* **7**: 3: 163–174.
- Torkkell, S., Ylihonko, K., Hakala, J., Skurnik, M., and Mäntsälä, P. 1997. Characterization of *Streptomyces nogalater* genes encoding enzymes involved in glycosylation steps in nogalamycin biosynthesis. *Mol. Gen. Genet.* **256**: 203–209.
- Ward, J.M., Janssen, G.R., Kieser, T., Bibb, M.J., Buttner, M.J., and Bibb, M.J. 1986. Construction and characterization of a series of multicopy promoter-probe plasmid vectors for *Streptomyces* using the aminoglycoside phosphotransferase from Tn5 as indicator. *Mol. Gen. Genet.* **203**: 468–478.
- Wiley, P.F., Kelly, R.B., Caron, E.L., Wiley, V.H., Johnson, J.H., MacKellar, F.A. and Mizsak, S.A. 1977. Structure of nogalamycin. *J. Am. Chem. Soc.* **99**: 542–549.
- Wiley, P.F., Elrod, D.W., and Marshall, V.P. 1978. Biosynthesis of the anthracycline antibiotics nogalamycin and steffimycin. *B. J. Org. Chem.* **43**: 3457–3461.
- Wright, F., and Bibb, M.J. 1992. Codon usage in the G+C -rich *Streptomyces* genome. *Gene* **113**: 55–65.
- Ylihonko, K., Hakala, J., Niemi, J., Lundell, J., and Mäntsälä, P. 1994. Isolation and characterization of aclacinomycin A-non-producing *Streptomyces galilaeus* (ATCC 31615) mutants. *Microbiol.* **140**: 1359–1365.
- Ylihonko, K., Hakala, J., and Mäntsälä, P. 1995. Process for producing anthracyclines and intermediates thereof. WO 96/10581.
- Ylihonko, K., Hakala, J., Kunnari, T., and Mäntsälä, P. 1996a. Production of hybrid anthracycline antibiotics by heterologous expression of *Streptomyces nogalater* nogalamycin biosynthesis genes. *Microbiol.* **142**: 1965–1972.
- Ylihonko, K., Tuikkanen, J., Jussila, S., Cong, L., and Mäntsälä, P. 1996b. A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater*: sequence analysis and complementation of early-block mutations in the anthracycline pathway. *Mol. Gen. Genet.* **251**: 113–120.
- Ylihonko, K., Hakala, J., and Kunnari, T. 1998. Hybrid anthracyclines from genetically engineered *Streptomyces galilaeus* strains. No. 981062.
- Yoshida, M., Fujioka, A., Nakano, K., Yuasa, C., Toko, T., Takeda, S., and Unemi, N. 1996. Activity of menogaril against various malignant lymphoma cell lines and a human lymphoma xenograft in mice. *Anticancer Res.* **16**: 2875–2879.

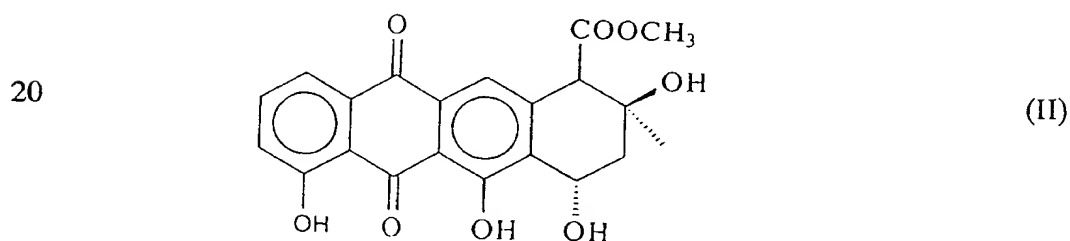
Claims

1. Isolated and purified DNA fragment, which is the gene cluster for the anthracycline biosynthetic pathway of the bacterium *Streptomyces nogalater*, being included in a 10kb
5 and a 7kb flanked *Bgl*II fragments of *S. nogalater* genome.
2. The DNA fragment according to claim 1, comprising the nucleotide sequence given in SEQ ID NO:1, or a sequence showing at least 80% homology to said sequence.
- 10 3. A recombinant DNA, which comprises the DNA fragment according to claim 1 or 2, cloned in a plasmid replicating in *Streptomyces*.
4. The recombinant DNA according to claim 3, which is the plasmid pSY15c, comprising a 1.4 kb *Bam*HI–*Sac*I fragment from the plasmid pSY42 and a 1.1 kb *Mlu*I–*Kpn*I
15 fragment from the plasmid pSY43.
5. Plasmid pSY42, deposited in *S. lividans* strain TK24/pSY42 with the deposition number DSM 12451.
- 20 6. Plasmid pSY43, deposited in *S. lividans* strain TK24/pSY43 with the deposition number DSM 12452.
7. A process for the production of hybrid compounds, comprising transferring the DNA fragment according to claim 1 or 2 into a *Streptomyces* host, cultivating the recombinant
25 strain obtained, and isolating the compounds produced.
8. The process according to claim 7, wherein the *Streptomyces* host is a *Streptomyces galilaeus* host.
- 30 9. The process according to claim 8, wherein the *Streptomyces galilaeus* host is selected from the strains H026, H039, H063 and H075, which are mutant strains of *S. galilaeus* ATCC 31615.

10. The process according to claim 8, wherein an anthracycline is produced, which has the following formula I



11. The process according to claim 8, wherein an anthracyclinone is produced, which has the following formula II



12. A process for the production of hybrid compounds, comprising transferring at least one of the genes selected from the group consisting of *snogJ*, *snogA*, *snoaM*, *snogN*, *snoaG*, *snogC*, *snogK*, *snoaL*, *snoK*, *snogD*, *snoW*, *snogE*, *snoL*, *snoO* and *snoaF* into a *Streptomyces* host, said genes being derived from the DNA fragment of claim 1 or 2, cultivating the recombinant strain obtained, and isolating the compounds produced.

13. The process according to claim 12, wherein the gene *snoaL* encoding NAME cyclase is transferred into a *Streptomyces* host.
14. The process according to claim 12, wherein at least one of the genes *snogD* and
5 *snogE* encoding glycosyl transferases is transferred into a *Streptomyces* host.
15. The process according to claim 12, wherein at least one of the genes *snogJ*, *snogN*, *snogC*, *snogK* and *snogA* affecting the formation of nogalamine and nogalose is transferred into a *Streptomyces* host.

1/4

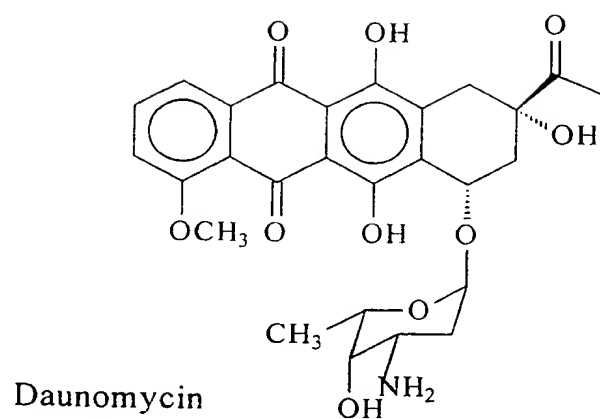
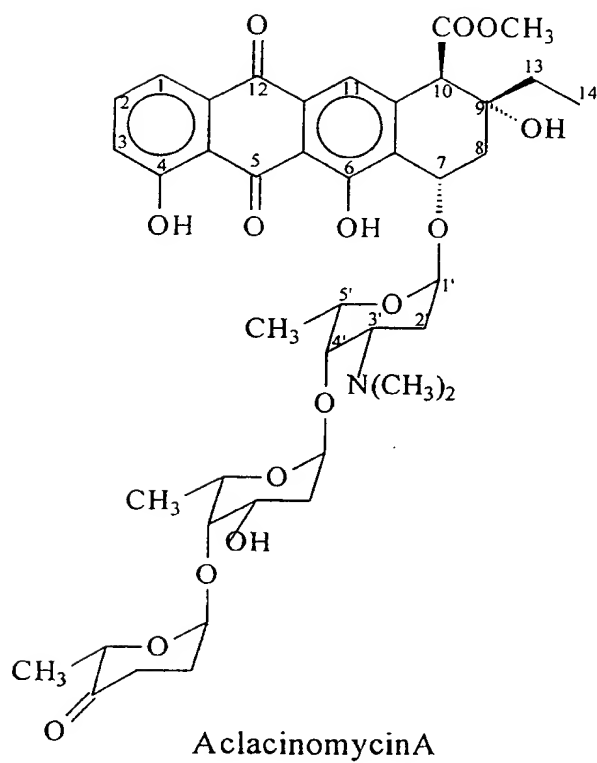
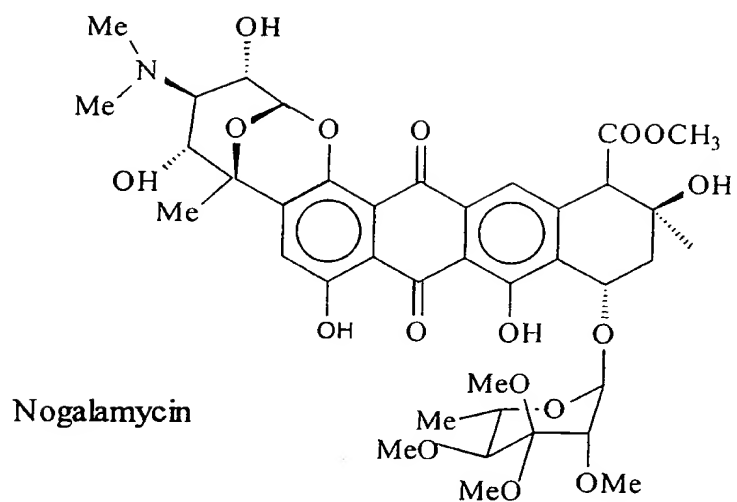


Fig. 1

2/4

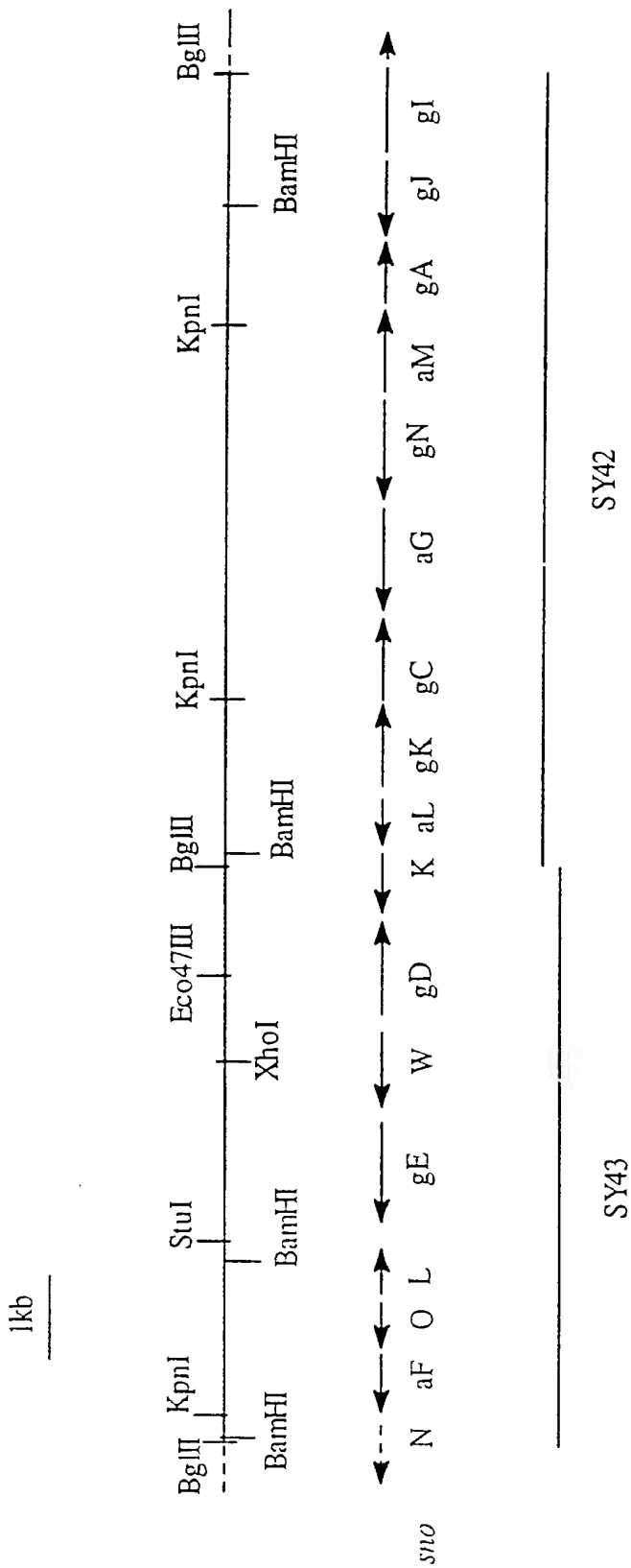


Fig. 2

3/4

Aglycone moiety pathway

Sugar residue pathway

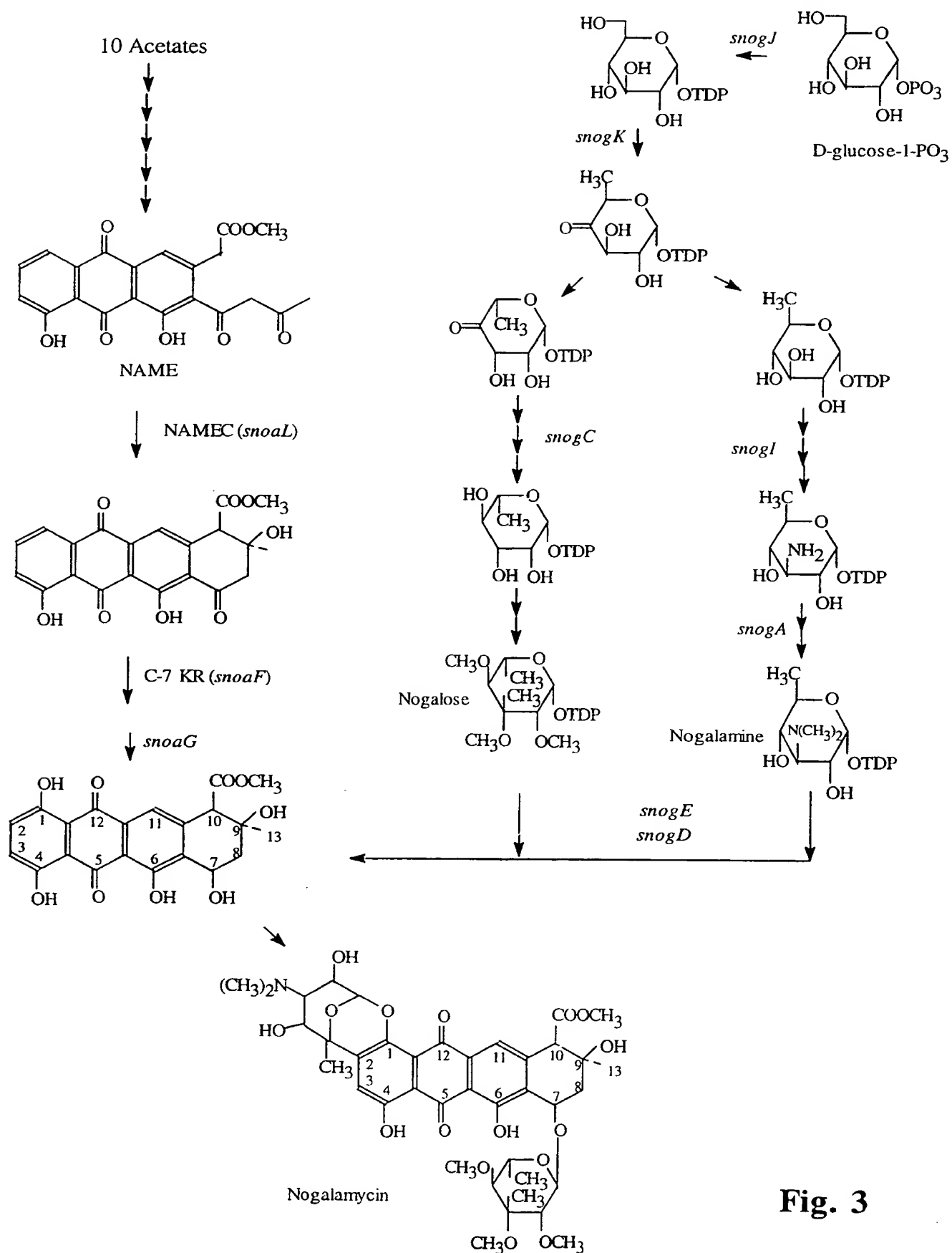


Fig. 3

4/4

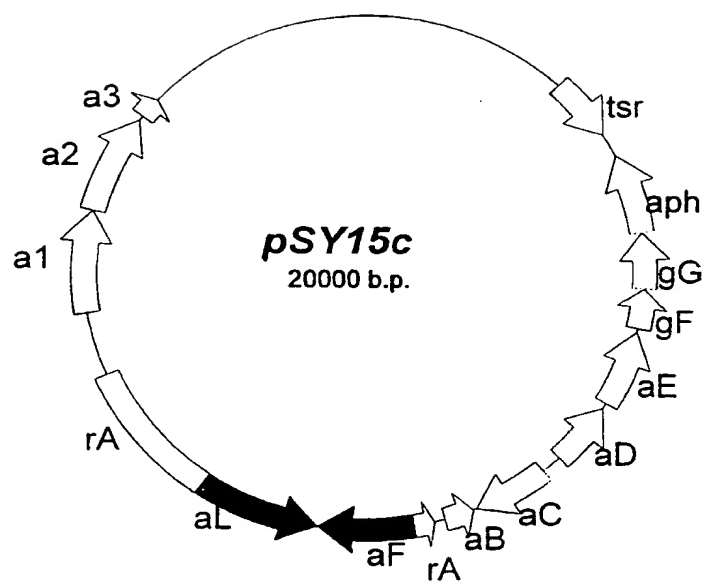


Fig. 4

SEQUENCE LISTING

<110> Galilaeus Oy
 <120> Gene cluster involved in nogalamycin biosynthesis, and its use in production of antibiotics
 <160> 18
 <210> 1
 <211> 16020
 <212> DNA
 <213> *Streptomyces nogalater* ATCC 27451
 <220>
 <221> misc_feature
 <222> 3799..3800
 <223> "overlapping sequence in the genes *snoaM* and *snogN*"
 <221> misc_feature
 <222> 6334..6356
 <223> "overlapping sequence in the genes *snoaG* and *snogC*"
 <221> misc_feature
 <222> 13201..13300
 <223> "unknown region"
 <400> 1
 agatctcgtc cgccagtgcc tcggtgaccg gcaacgagcc cttggcgtag ccgagatggg 60
 agaaaccggg catggtgtgc acgggccagg gataactgat gttgagggcg atgtcgtagg 120
 aggcgcgcag ggcctccagc accgcgtccc gtcgcggatg ggcaccacg tacacgtagt 180
 agacgtgctc gttgccctgc gcggtcctcg gcagcagcag ccccggtgtcc gccaggccct 240
 cctcatagcg gcgtgccacc gcccgggcgg cctcgatgta ggacggcaac cgggacagct 300
 tgcgcgcgag gatctctgcc tgtacttcgt ccagccgggt gttgtgcccg ggggtttcga 360
 cgacgtagta gcggtctctc atgccgtagt agcgcagccg ccgcagccgg tccgccaccc 420
 gctcgtcgtc ggtgagcacc gcgccgccgt ccccgtaacg gcccagcacc ttggtcgggt 480
 agaaggagaa cgcggccgcy tcaccggtcg agccggcgag tcggccgtgc cggcgcgcgc 540
 cgtgcgcctg cgcgcagtcc tccaggatca ccaggttgtg ccgggcggcc agatcgcgca 600
 gcggtgccat gtccacgcac tgcccgtaga ggtggaccgg cagcagacac cgggtgcgtg 660
 gcgtgaggac ggccctccacc tgggacgtgt ccatcaggtg gtcctcctcg cgcacgtcca 720
 cgaagacggg cgtggcaccg gccgagtcga tcgcgacgac cgtgggcgcg gcggtgttgg 780
 acacgggtgac gacctcgtcg ccggggccga caccgaaggc ctgtaacccc agcttgacgg 840
 cggttggtccc gttgtcgacg ccgacggcat gtccgacgcc ctggaatgag gcgaactcgg 900
 actcgaagcc gcgcacgctc tcaccgagga cgagccggcc ggagcggaac accgtctcca 960
 cggcatcgtg gatgtcctcg cgttccagct cgtattccgg cagatagtcc cacacgtgta 1020
 cggtcacgca gccctccgg gattctccct gcgaaaagtc accactctac gacaacgttc 1080
 accactcgct ttttcctcaa cgtccgcttg agaaggcccg gcctgctgtg gcccggggaa 1140

aggtgcggtc	ggtatcatcg	actccgtctt	ctcattcggga	ggttggttcag	ggtgaaggga	1200
atcattctcgc	ccgggggtac	ggggagcagg	ctccaccgga	cgactctcgc	ggtgtccaag	1260
cagcttctcc	ccgtcgggga	caagccgatg	atctactacc	cgctctccgt	gctgatgctg	1320
gccggcgtca	cggacatcct	catcatcagc	acaccgcacg	aactcccccg	aatgcgccgt	1380
ctgttcggcg	acggcgcaca	gctcggactc	cgccctggcct	acgccgagca	ggagaaaccc	1440
aggggtatcg	ccgaggcggt	cctgatcggt	gccgaccacg	tgggaagcga	tgccgttgcg	1500
ctggcgctgg	gcgacaacat	attccacggg	agttcttttc	agggggtgct	gcgcaaggaa	1560
gccgaggaat	tggacgggtg	tgtcctgttc	ggttatccgg	tcaaggatcc	ccagcgttat	1620
ggagtcggcg	aggcgaacgc	gtccggggcg	ctcgtctcca	tcgaggagaa	accggtacgc	1680
ccccgctcca	accggggccat	caccggactc	tatttctacg	acaacgaggt	ggtggacatc	1740
gcccgggcggc	tgcgcccctc	cgcccgcggc	gaactcgaac	tcaccgacat	caaccgtacc	1800
tacatggaac	gaggccgggc	ccggctcgtg	gacctggggc	ggggattcgc	ctggctcgac	1860
accggcacac	ccgagtcact	cctgcaggcc	tcgcagtacg	tgtccgccct	ggaggaacgc	1920
cagggcacat	ggatcgccct	catcgaggag	gtggccctcc	gcattgggctt	catcaacgcc	1980
caggcctgct	acgaactggg	cgcgcgccct	tccggctccg	gctacgggca	gtacgtgatg	2040
gccatcgccg	aggagtgcac	ggggcggggt	tgagcgggcg	tgccgggtgg	gcgaacggcc	2100
cggccttacc	cgcccccgcg	caccccgacg	aacaaccccc	ggccgggtcag	cccgtcgtcc	2160
aggaactcgg	ccgggcagcc	cgcgctcctc	aacgcggcga	ggtactcctc	cctgggtgaac	2220
agggtgagca	ggtcgatctc	cgtgaactcg	cgtatcccgg	tggcctcgcc	gaccaggaac	2280
cgcacctcca	tgcgggtcct	gcggccctgc	ctggtggagt	gggacaccgg	ggccacggtc	2340
cggccttcac	cgcgtgccag	gtccccggcg	acgtagccct	ccaggaaccg	ctcggggaac	2400
caccagggct	ccaccacgag	cacgccgccc	ggcaccaggt	gcgcggccat	cgtgcgcacc	2460
gccgcccgcg	tgtccgcgac	ggtctccaga	taccgatggg	agcagaacag	gcagaccacg	2520
gcgtcgaaac	gcccgcctcag	ggcgaagtcg	cgcattgtccc	cgggccgcac	cggcaccccc	2580
ggcagccgcc	gttcggccag	ggcccgcatc	tcgtccgaca	gctccaggcc	ctccgtgtgc	2640
gcgaacagcc	cgcggaaggc	ctccagatgg	gcgccgggtg	cgcaggcgac	gtcgagcagc	2700
gaacgcgccc	cgggccgacg	ggacctgatc	tccgcgggtg	ccggttcggc	ctcgtccgcc	2760
cagctctttc	cccggctcgc	gtagaccatc	tcgtacacgt	ccgccagttc	ccggccgtac	2820
acgcgtcagt	cctcgtccac	cagggcgacc	gcccgggtcc	acccggcgcc	ggcgccggcg	2880
accttgaccg	ggaagcagca	gacgcggaac	ccgaaggaga	ccggcaggcg	gtcgaggttc	2940
gccagccgct	cgatctggca	gtactcccgc	tcccggccca	ccacgtgcgc	gggccacagc	3000
accgatcggt	cgccggctgc	gcggtaccgg	tcgatgatgt	ggccgaaggg	cgcgtccagg	3060
ctgaaggcat	cggctccgat	cacccggacc	ccgtgggtcg	gaagcatccg	taccgcgggc	3120
ccgtcgagac	cggcgaagtc	cgtgaagtag	cgcgggggtg	ccgcgtgccg	ctgggcaccg	3180

gtgtgcagca	gcaogatgtc	cccgggccgc	aacgcgcacc	cggtcggggc	cagttccttc	3240
tccaggcgcg	cggcgctcac	ggtgcccgtc	ggagcgtcgg	tgagggtccag	caccaccccg	3300
cgcccgaaga	accactccag	cggcatctgg	tcgatgtggc	gggggacgcc	gtccccgtac	3360
agcgcgcgcg	aaccatagtg	cgacggcgcg	tcgacgtgcg	tgccgggtgtg	cgtgggtcagc	3420
gtgatcctgt	ccagtgcacg	gaactcgccg	tccggcagtt	cgtccggaga	gaactcgaca	3480
ccgaagtgtc	cgcgcacatc	cgcgcacatg	tggtccgcgc	cctgccgggg	cgtgaggacg	3540
tcgtgcacca	ccgggtcggg	ctcgtactgt	gaggaatcca	ccggtgacga	aaggtcgatg	3600
agccgcacgc	gcacctccgg	gttcgtagac	gggtctcggt	gacgcagcgc	gggtacgacg	3660
ctgacacgcc	cctcttgacg	tggcctggaa	gctggttcga	cgggcgggca	ccgcacgcga	3720
cggccggcgc	cgcaccggcg	ccgtcccggc	cgagcgggaa	tccagggagg	gtatagcggc	3780
gcgccccacg	ctgccgtcat	ggtgatgaaa	ctgacggaca	gcgagctggg	gcgtgcgctg	3840
ctctcgctgc	gtggttacca	gtggctccgc	ggcatccacc	acgatcccta	cgccctgctg	3900
ctgcgcgccg	agagcgacga	tccggcgacg	ctcggccggc	tgctgcgtga	acgcggccgg	3960
ctccaccgca	gcgacaccgg	cacctgggtc	accgcggacc	atgcgacggc	ctcccggctg	4020
ctcgccgacc	cgcgcttcgt	gctgcgcgcg	ccgccggccg	ggcccggcac	cggcaccggg	4080
gacgtcatgc	cgtgggaaga	ggccacgctg	agcgacctgc	tgcccctcga	cgaggcgcg	4140
ctgacgaccg	accgggcacg	gtgccgcggg	ctcggcgcgga	ccgccgcgcg	gatcgcgggc	4200
gacggtcccc	tcgcgacgcg	actcgcgacg	ctggccgggg	cccagaccga	acaggtgcgc	4260
tcaacggggc	acttcgacct	cagggccgac	tacgccctcc	cgtacgcggg	cgagccggcc	4320
tgcgcgctgc	tcggcctgcc	ggccgggcag	tggtccctct	tcggcgccct	ctccccggcc	4380
gtcctgctcg	acgcgacggg	cgtaccgccc	cgccttcggg	aggcgcgcg	cctgatcgcc	4440
tccacggcgg	aactgaccgc	cctctggccg	cggctggccc	cgagcctgtc	gaagaccgtc	4500
ccggaggacg	aagcgccgga	cctcttcctg	ctgacggccg	tggtactcgt	accggccgtc	4560
gtccacctgg	tctgcgaggg	ggtcgccgcc	ctgtcgacacg	accccgggca	ggccgggctg	4620
ctcagggacg	acccggtact	cgcgcacacc	gcggtcgagg	agacgctgcg	ccacgcaccg	4680
cccgcgccgc	tggtcaccc	ccacgcgacc	ggaccggagc	gcgtcgcgga	cgtcgacctc	4740
cccgcggggc	ccgaggtcgc	cgtcgctcgt	gcggcgggcg	accgcgatcc	ctcctgggtg	4800
ccggaccccc	accgcttcga	cctcaccagg	aacgagcggc	atctggcaact	gccgccggat	4860
ctgccgctgg	gggcgctcgc	ccgcgtgctg	cgcgtctgcg	cgaccgcggc	cgtcgcgggc	4920
ctcgcgggcg	gactcctccc	gctgcggggc	gtcgggccgc	ccgtacgacg	gctgcgtgcc	4980
ccggtcacc	ggtccgtgct	gcgcttcccc	gtcgccccgt	gctgagcagc	ccctcctcac	5040
gtcatccccc	gcccgccttc	ccccgcccgc	aacggaaggg	actctccatg	gacaaccgcg	5100
agaccgtacg	accggtgagc	gtctgccggg	tctgcggcgg	caacgactgg	caggacgtcg	5160
tggacttcgg	tgacgttccc	ctcgccaacg	gcttcctgtc	cccggccgac	tcctacgaga	5220

acgagcgccg ctacccgctg ggcgtcctgt cctgccgcgc ctgccggctg atgagcctga 5280
cccacgtggt cgaccccgag gtgctgtacc gcgactacgc ctacaccacc cccgactccg 5340
aaatgatcac ccagcacatg cggcacatca ccgcgctgtg ccgcaccctt ttcgagcttc 5400
ccccggacag cctcgtcgtg gagctgggca gcaataccgg ccgtcagctc atggccttcc 5460
gcgaagcggg gatgcgccacc ctgggcgtgg accccgcgcg caacctcacg gacgtcgccc 5520
ggcgcaacgg catcgagacc ttccccgact tcttctccca cgacgtggcc cgcaccatcc 5580
ggcgcgacca cgggcaggcg cggctcgtgc tgggacggca tgtcttcgcc cacatcgacg 5640
acgtgtcgga catcgcgccg ggcgtacgcg aactcctgtc tcccgacggg gtgttcgga 5700
tcgaggtgcc gtacgttctg gacctgctgg agaaggtcgc gttcgacacc atctaccacg 5760
agcacttgtc gtacttcacc atgcggtcct tcgtcaccct cttcgcgcg caccgggtgc 5820
gggtgctcga cgtggagcgg ttccggcgtgc acggcggtac ggtcctcgtc ttcgtgggccc 5880
acgaggacgg cccttgggcc gaacgtccct ccgtccccga actgctgcgc gtggaacggc 5940
agcggggcct ctacgacgac gccacctacc gcacgttcgc gcagcggatc gagcgggtgc 6000
gcaccgaact gccggaactg ctgcgctccc tcgtggccca gggcaagcgc atcgtcggct 6060
acggtgctcc ggccaagggc aacaccatcc tcacggtgtg cgggctcggc ctgaaggagc 6120
tggaatactg caccgacacc accgagctga agcagggcag ggtgctgccc ggcaccaca 6180
taccggtgca cgctcccgag caccgcaagg aacacatccc cgactactac ctgttgctcg 6240
cctggaacta cgccacggag atcctcgaca aggagacggc cttccgggac aacggcgggc 6300
ggttcacgtg gcccatcccc cgcccgctga tcctcacgtc ccgctcaggt tcctgaggcg 6360
cccgccgggc agcagctgac gcacgcctc gcgcagggt gcacgccagt cgcggggcgg 6420
tgcgacgccg accagccgcc agcggctcgtg cccgagcacc gtgcacgccg gccggggcgc 6480
cgggcccggc cggtcgggcc tgcgccaccg gcgcaccctg tccgggtccg cggccgccag 6540
ccggaacacc tcccgggcca gctcgtacca ggtggccgcc ccggcggttg tggcgtggaa 6600
gatcccgcgc gcccggtctg gcggcggtgc gccagcgtc accagcagcc gggccacgtc 6660
accggccccac gtcggctgcc ccactggtc gttgacgacg tcgacatggc cgtcgtccgg 6720
ggcacgctcc agcatcgtgc gcacgaagct gcggccctgc ccgccgtaga gccacgccgt 6780
gcgcaccacg gtgcccgtat ccggcagcag cgacagcacg gcccgttccc cggccagttt 6840
gctgcggccg tacaccgtgc gcgggcccgg agcgtccgac tcgccgtaag ggctgcgggt 6900
gtcggccggg aagacgtagt cggtcgagac gtggatcagc cgtacgccgt ggcgcgcaca 6960
gcggcgggcc agcagccggg gcccgccgcc gttgacgcgc atgcctccg cccaccgca 7020
ctcggcgccg tccacgtccg tgaaggcggc gcagttgacc accaccgcg gccgggtgcgc 7080
ggcgaacgcg gcgtccaccg cccgcccgtc ggtgatgtcc agcgcgcgcc gcccgagtac 7140
caccgcctcg gcggcgggcc ggctcctgcc ggtctccgcc agggccgcgg tcaggtgccg 7200
ggcgagcatg ccttctcctc cggtgaccag cagcgcacac ccgctcaccg gaccccgggg 7260

acgacggtgg	acgtaccgcc	cggcgcctgt	actccccgct	tgagcggctc	ccaccaggac	7320
cggttctcgc	ggtaccactg	gaccgtcgag	cgcagccccg	aggagaactc	ccgcgccgga	7380
cggtagccca	gttcctcacg	ggccctgccc	cagtccaggc	tgtaacgcag	gtcgtgcccc	7440
ttgcggtcgg	gcacgtgccg	gacgctgctc	cagtccgccc	cgcacagctc	cagcaacata	7500
cccaccagct	cccggttgga	gagctcccgg	ccgccgccga	tgtggtacac	accgccgggc	7560
cggcccgcgg	tgccgaccag	gtccacgccc	cggcagtggt	cctccacgtg	cagccactcc	7620
cgcacgttcc	gcccgtcccc	gtacagcggc	accggcagcc	cgtccaacaa	gttggtgacg	7680
aagcgcggga	tgagcttctc	cgggtgctga	cgcggggcgt	agttgttgga	acagcgggtc	7740
acccgcacgt	ccaggccgtg	cgtgcggtgg	caggcgaacg	ccatcaggtc	ggccgacgcc	7800
ttggaggcgg	cgtacgggga	gttggggctc	agcgggtgct	cctccggcca	ggaaccggac	7860
gcgatggagc	cgtagacctc	gtccgtggac	accaggacga	agggtccac	gccgtggcgc	7920
agcgcggcgt	ccagcagccg	ctgggtgccg	acgacgttgg	tcagcacgaa	gtcgtcggcc	7980
gcgcggatgg	accggtcgac	gtgcgactcc	gcggcgaaat	ggacgacctg	gtcgtgtgt	8040
gccatcagct	cgtcgaccag	ctcggcgtcg	aggatgtcgc	ccgcacgaa	gcgcagccgg	8100
tcaccgcgta	ccgcgtccag	gttcgtgagg	ttgcccgcgt	acgtcagttt	gtcgaggacg	8160
gtgacgcgta	ccgcgcgggc	ccccgctccg	ggggcccggg	tctccagcag	catgcgcaca	8220
taggccgagc	cgatgaaacc	gaccgcgccg	gtgaccagga	tgttcacgtc	cgctcgtcgcg	8280
gaggtgtgcg	acgccatggg	ttccctccat	ccgtcgggtg	ccgtggggcg	gagtgcgccc	8340
cctcgaccca	gcgtcggggg	cggccgtgga	ggagcggttg	agcttcggcg	cagcggcggc	8400
tcgaccggcg	gcggccggcg	tcgccggact	ccaacggttc	tcgacggaac	gaccaacggc	8460
cctggcgaga	ctgcccggac	agcccggccg	agagaggag	gacccttga	gccgtcagac	8520
agagatcgtc	cgcccgatgg	tgagcgcctt	caacaccggc	aggaccgacg	acgtggacga	8580
gtacatccac	cccgaactac	tcaatccggc	caccttgga	cacggcatcc	acaccgggcc	8640
caaggcgttc	gcccagctgg	tcggctgggt	gcgggcgacg	ttctccgagg	aagcccgcct	8700
ggaggaggtg	cggatcgagg	agcgcggccc	gtgggtcaag	gcctacctcg	tgctctacgg	8760
ccgccacgtc	ggccggcttg	tcggtatgcc	gccaccgac	cggcgcttct	ccggtgaaca	8820
ggtgcacctg	atgcgcatcg	tcgacgggaa	gatccgcgac	caccgggact	ggcccgaactt	8880
ccaggggacg	ctgcgccagc	tcggcgaccc	gtggcccgcg	gacgagggct	ggcgtccgtg	8940
accgtccctg	aaaccgcacc	cgacgagaca	tcgaccagg	aaggatggct	catgccggat	9000
cccggcggcc	cgaccacggc	cgagaacctg	tcgaaggagg	ctgtccgctt	ctaccgcgag	9060
cagggttacg	tgacatccc	gcgcgtcctg	tcggagacgg	aggtgaccgc	cttcggggcc	9120
gcctgtgagg	aggtcctgga	gaaggagggc	cgcgagatct	ccggcatcgc	cctgcggctg	9180
gccggcgcg	ccctgcgggt	ctacagcagc	gacatcctgg	tcaaggagcc	caagcgcacc	9240
ctgcccaccc	tggtccacga	cgacgagacg	ggactgccgc	tgaacgagct	gagtgccacg	9300

ctgacggcct	ggatcgcgct	gacggacgta	cccgtcgaac	gcggtgcat	gagctacgtg	9360
ccgggctccc	atctcagggc	ccgcgaggac	cggcaggagc	acatgaccag	cttcgccgag	9420
ttccgggacc	tcgcggacgt	gtggcccgat	taccctggc	agccgcgcgt	cgccgtgccc	9480
gtccgcgcgc	gagacgtcgt	gttccaccat	tgccgtaccg	tccacatggc	cgaagccaac	9540
accagcgact	cggctccgcat	ggcgcacggc	gtcgtctaca	tggacgcgga	cgccacctac	9600
cggccggggc	tccaggacgg	ccacctgtcc	cgcctgtcgc	cgggagatcc	actcgaaggc	9660
gagctgttcc	ccctggtcac	ggcaggcaca	cggcagttag	gtccgcgcgt	cccggcggtc	9720
gcgggacgcg	cggggacggc	accgtcagcc	ggccagcgcc	acgagcttgg	cggccgtctc	9780
ggccggcggc	ggcatctcgc	tcatctcctg	ccgcacccgc	agggccgcct	cccgcaaccc	9840
cgcgtcgtcc	agcagccgtc	ggcactgtc	ggcaccagc	gatcccgcc	cggcatcgaa	9900
cccgatgccc	agcccgggtca	gcacatcgcg	gttgggtgtcc	tggtaggagc	cgtgcgggat	9960
gacgcactgc	gggacgccgg	cggccagggc	cgtcagcagt	gtgccgctgc	ccccgtgatg	10020
gatgatcgcg	tcgcacgtct	ccagcagcgc	gccagcgga	atccactcca	ccaccgggtac	10080
gttcgcgggc	agttcaccga	gcagggccag	gtcgcgcgcg	cccagggtca	gcacgaactc	10140
cgcgtccacg	tccgccactt	cggagaacag	cggggccagc	ttggcgatgc	cgcccacag	10200
cgcgtcgatg	gagcccagcg	tcaccgcgat	acgccgccgg	ccggccgcgc	gcggcagcca	10260
gtccggcagc	accgtctcgc	cgttgtaggg	gacgtaccgc	atcggccagg	caccggggga	10320
gcgccgggtcc	tccggcagca	gcgcctccac	gctcggcggt	gtcgtcgtca	gccgcacgga	10380
accggtcggc	tcgccggtga	cgccgtggcg	ctcgtagtcc	ttggacatcg	cccgccggat	10440
gagcgcgcgc	agccccggct	cgtgttccgc	gggaccagc	ggcagctcta	cgacggcgag	10500
ttgcagcgct	gccgccgtca	gcggggccgc	gccctgtgtc	ggagtgtgca	cgacgaggtc	10560
gggccgccag	ctccgcgcgc	tccgcagcgc	cccgtcgacg	gccaccgccg	ataccggggc	10620
gaacatctcg	gcgaagaagc	cctcgcccag	cccctcggag	tgcacgggt	cggtgacgtc	10680
ggtgtcgtcg	ggcacgaaca	gcttcgcgta	gttcacgcgc	ggcgacacgt	ccacggcgca	10740
cagcccggcc	tccgcgacgg	cgcggatgtc	gcccccgctg	gcgtagcgga	cctcgtggcc	10800
gagagcgcgc	agcgcctgtg	ccagcggcac	cgtcggcagg	atgtggctga	gcccggtga	10860
agtgatgaac	aacgcacgca	tgatgcccc	tggtcgacat	gaacctggaa	cacgcaccc	10920
gacggcgcct	tctgttgctc	cggtcgacgc	ccggtcgaca	ggccctcgta	cagcccgcgc	10980
ggggccgggtc	cggccacgac	gcaggctcca	gcggacgtcg	acggcgggga	cgcagcgtgg	11040
tcgccggggag	gcatcgatga	cagtattggt	aaccggagcc	acaggaaacg	tcggccggca	11100
cgtcgtcacc	gggctactgg	ccgccggccg	ccgggtgcgc	gcgtgaccc	gcacaccgga	11160
ccggtccggc	ctgcccggcg	gcgcggagat	cacaggcggc	gacctgaccc	gcccgagac	11220
ctacgagcgc	atgctggacg	gtgtcgaagc	cgtctacctg	ttccccgtcc	cggagaccgc	11280
cgcggcgctc	gccggggccg	cgcgacgggc	cgggtgtccg	cggatcgtgg	tgctctcctc	11340

ggactccgtc accgacggca ccgacaccgg aggcacaccgg cgcgtggaac tggccgtgga	11400
ggacacgggg ctcgagtgga cccatgtgcg ccccggcgag ttcgcgctca acaaggtcac	11460
cctgtgggcg ccgtcgatcc gcgcggaggg cgtcgctcgg tccgcgtatc cggacgcccg	11520
ggtggccccc gtgcacgagg ccgacgtcgc ggccgtcgcg gtgaccgcgc tgctgaagga	11580
ggggcacgcc ggccgcgcct acagcgtgac cggaccgcag gccctcacc agcgcgaaca	11640
ggtccgcgcg gtaggggagg ggctcggccg gtccctcgcg ttcgctcgagg tgaccccccg	11700
gcaggcgcgg gccgacctga ccgcccaggg gctgcccgcg cccatcgccg actacgtcct	11760
cgccttccaa gccgggtgga ccgagcggcc cgcggccgcg cggccgaccg tcggggaggt	11820
caccggccgg cccgcccgcg cgctcgccca gtgggcccgc gaccaccgag cggacttccg	11880
gtgaccggag accgcgtcca ccgcgccacg acagaaaggc gacgcccggt cgcgtactgc	11940
tgacgtcctt cgccatggac gccacttct gcaccgccgt gccgctggcg tgggcaactgc	12000
ggtcggcccg gcacgaggta cgggtggccg gccagcccgc gctcacctcc accatcacgg	12060
gagccggcct gaccgccgtg ccggtcggcc gcgaccacac gcacggcagc ctctggggcc	12120
gggtcggcag cgacatctc gccctgcacg acgaggcgga ctacctggag gcccgtcacg	12180
acgccctggg ctctgagttc ctcaaagggc acaacacggg gatgtccgcg ttgttctact	12240
cgcagatcaa caacgactcg atggctcgac acctggtgga cttcgcccgt cactggcggc	12300
ccgacctggt cgtctgggag ccgttcacct tcgcgggcgc cgtggccgcg cgggcctcgg	12360
gcgcgcacca cgccgcctg ctgtccttcc ccgacctgtt cctcagcacg cggcgctct	12420
tcctggagcg catggcgcg ccaggagccc agcatcacga cgacacactc gccgaatggc	12480
tcgactggac ccttggcccg cagggccact ccttcgacga ggagatcgtc acggggcagt	12540
ggtccatcga ccagaccccc gcccccgtgc ggctcgacgc cggcgggtccc accgtgccga	12600
tgccgttacgt cccctacagc ggactggtgc ccacagtggg gcccgactgg ctgcgcaggc	12660
cgcccgagcg gccacgggtc ctggtcaccc tcggcatcac ctacggcgcg gtgaagtcct	12720
tcctcgccgt ctccgtggac gaccttttcg aggcctgggc cgggctcggc gtcgaggtgg	12780
tcgccaccct cgacgccgac cagcgggagc tgctggggcg cgtgccggac cacttccgca	12840
tcgtcgagca cgtgccgctg gacgccgttc tgccgacctg ctggcgatc gtccaccacg	12900
gcggagccgg cacctggtcg acggccgcgc tgtacggggg gccgcaggtc tccctgggct	12960
cgatgtggga ccacttctac cgggcccgtc gcctggagga actcggggcg gggctgcggc	13020
tgccctccgg cgagctgact gccgaggggc tgccgacccg gctggagagg gtgctcggcg	13080
agccctcctt cggcaccgcc gcgcaggcgc tgagcgacac catcgcggcg gaaccagcc	13140
ccagcgaggt cgtgccggtc ctggaggagc tgaccggacg gcaccgtccc ggcaccggg	13200
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	13260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn ccgtccgggc ccctcgccgg	13320
tgagggagcc cggatcacag tccgtccggc accacgcccc ggtcccggaa cagcggggag	13380

aagttgaaga	cgtcccagtg	ctccacgacc	ttgccggctt	cggagaagcg	cagctcctcc	13440
aagtaggtcc	agcggacctt	gcggccggtg	ggggcgatgc	ccatgaacac	gccctggtgc	13500
gtggccgagc	aggtgatccg	cagcatcacg	cggtcgccct	cgcccacgat	gctccgcacg	13560
tccagacgaa	ggtccgggaa	ggcctccacc	gcgctgttca	tacgccgtac	gacctcctcg	13620
gcgctcaccg	gtttgtcctc	gtcgtcgtag	tggacgacgt	cgggtgcccc	gtgcgcgacc	13680
accccggaga	cgtcccaccg	gttccatgcg	gccaccatct	ccaggcagcg	ttccttgttc	13740
gcggtcgttg	acatgtcgac	tccttgaagg	cccgggacta	ctggtcacgc	gccagccttc	13800
caacccgccc	cggaaaagcg	gtgcacgacc	gctggagccc	gcaccggaac	ctgcgcggcg	13860
gagctgaacg	gggtttcgag	ccgttcacca	aggacctgcc	gcagcctggt	acggcacacc	13920
ctgacgcctc	gctccgcgcg	ggacgcgccc	gccgggagga	aggacacacc	accatgtcgg	13980
tacgcaccga	tcagacggcg	gcaccggaag	accgagcggc	ggccacggat	cccgggttcg	14040
ggcacctgta	cgcgaggtg	cagcagttct	acgcccggca	gatgcagctc	ctcgactccg	14100
gcgcggccga	ggagtgggccc	gccaccttca	ccgaggacgg	cacgttcgcc	cggccctcct	14160
cgccggaacc	ggcacgcggc	cacgccgaac	tggccgccgg	cgcccgcgcc	gccgccgaac	14220
gcctcgccgc	cgagggcctt	tcgcaccggc	acgtcatcgg	catgaccgcg	gtacgccggg	14280
aacccgacgg	cagcgtgttc	gtacgcagct	acgccaggt	cttcgccacc	cgccgcgggg	14340
aagctccccg	gctgcactctg	atctgcgtct	gcgaggacgt	gctcgtgcgg	gaggggcccgg	14400
ggctgaaggt	gcgggaacgg	gttgtcacgc	acgacgcgtg	agggcggtcg	acccgccggc	14460
cgagccgcac	ctctgccacc	ccctcggcac	gccagccggc	gtcgagtccg	ctgcgagagg	14520
gcgcacttag	cgtgcgagcc	atgactgact	cgacaggtcc	ccgcccgggtg	cccgccatgt	14580
cacccgcccc	cagccccacg	ccttcccccg	gccccgcccc	cgggagcgaa	cccgcgcgcg	14640
tcgccgtgat	cgtcacccgg	ggcggttcgg	gtatcgggcc	ggccaccgcc	cgcgcccttcg	14700
ccgctcaggg	tgcgaaggtg	ctcgtcgtcg	gccgtaccga	ggacgcgctc	gcgcagaccg	14760
ccgagggctg	tgcggacatg	cgtgtgctcg	tcgccgacgt	ggcctcgccc	gacgggcccgc	14820
aggcggtcgt	caacgccgcc	ctgcgggagt	tcgggaggat	cgacgtcctg	gtcaacaacg	14880
ctgccgtggc	gggcatggag	accctgcaga	ccgtcgaccg	ggacgccgtg	gcacggcagt	14940
tcggcaccaa	tctgacggct	ccctctcttc	tcgtccagtc	cgcactcggc	gcgctggaga	15000
agtcgcgcgg	catcgtcgtc	aacgtgggga	ccgccgcgac	cctgggcctg	cgcgccgccc	15060
cgaccggcgc	gctgtacggg	gcgagcaagg	tggccctcga	ctacctgacc	cggacctggg	15120
ccgtcgaact	ggcccccccg	ggcatccgtg	tcgtcggcgt	ggcaccgggg	gtgatcgaca	15180
cgggcatcgg	cgtccgcgatg	ggcatgacct	cggagggcta	ccgggagttc	ctgaccggca	15240
tgggcccag	ggtgcccggtg	ggccgggtcg	gccgtccgga	ggacgtggcc	tgggtggatcg	15300
tccagctcgc	ccgcccggag	gccggctacg	cgacgggcat	ggtcgtcccc	gtcgacggcg	15360
ggctgtcgct	ggtctgaccg	gacaaggaag	gaaataccgc	aggaaggaag	taccgcagca	15420

```

aggaaatacc gcaggaagga gatatcgccg tgcaggaaac cgaacccggc gtccccgcgg 15480
acctgccccg cgagagcgac cctgccgccc tggagcgcct cgccgcacgg tacgggcggg 15540
acggctacgt ccacgtcccc ggcgtcctcg acgccgggga ggtcgccgaa tacctggccg 15600
aggcccgtcg gtcctcgcgc caccgaggagt ccgtgcgctg gggctccggc gccggcaccg 15660
tcatggacta cgtcgcccgc gccagctcgc gcagcgacac gatgcgccc cttgccaccc 15720
acccgcgcac cgccgcacct gccgagtacc tggccggctc gccctgagg ctgttcaagc 15780
tggaggtgct gctcaaggag aacaaggaga aggacgcctc ggtccccacc gccccgcacc 15840
acgatgcgtt cgccttcccc ttctccaccg ccggcaccgc cctgacggcg tgggtcgcgc 15900
tggtcgacgt cccggtggaa cgcggtgca tgacctcgt ccccgatca cacctgctgc 15960
cggatcccga taccggcgac gagccgtggg ccggggcctt caccggccg ggagagatct 16020

```

```

<210>      2
<211>      342
<212>      PRT
<213>      Streptomyces nogalater ATCC 27451

```

```

<220>
<223>      "translate of snogI, function: aminotransferase"

```

```

<400>      2

```

```

Met Thr Val His Val Trp Asp Tyr Leu Pro Glu Tyr Glu Leu Glu Arg
 1              5              10              15
Glu Asp Ile His Asp Ala Val Glu Thr Val Phe Arg Ser Gly Arg Leu
              20              25              30
Val Leu Gly Glu Ser Val Arg Gly Phe Glu Ser Glu Phe Ala Ser Phe
              35              40              45
Gln Gly Val Gly His Ala Val Gly Val Asp Asn Gly Thr Asn Ala Val
              50              55              60
Lys Leu Gly Leu Gln Ala Leu Gly Val Gly Pro Gly Asp Glu Val Val
              65              70              75              80
Thr Val Ser Asn Thr Ala Ala Pro Thr Val Val Ala Ile Asp Ser Ala
              85              90              95
Gly Ala Thr Pro Val Phe Val Asp Val Arg Glu Glu Asp Tyr Leu Met
              100              105              110
Asp Thr Ser Gln Val Glu Ala Val Leu Thr Pro Arg Thr Arg Cys Leu
              115              120              125
Leu Pro Val His Leu Tyr Gly Gln Cys Val Asp Met Ala Pro Leu Arg
              130              135              140
Asp Leu Ala Ala Arg His Asn Leu Val Ile Leu Glu Asp Cys Ala Gln
              145              150              155              160
Ala His Gly Ala Arg Arg His Gly Arg Leu Ala Gly Ser Thr Gly Asp
              165              170              175
Ala Ala Ala Phe Ser Phe Tyr Pro Thr Lys Val Leu Gly Ala Tyr Gly
              180              185              190

```

10

Asp Gly Gly Ala Val Leu Thr Asp Asp Glu Arg Val Ala Asp Arg Leu
 195 200 205
 Arg Arg Leu Arg Tyr Tyr Gly Met Glu Ser Arg Tyr Tyr Val Val Glu
 210 215 220
 Thr Pro Gly His Asn Ser Arg Leu Asp Glu Val Gln Ala Glu Ile Leu
 225 230 235 240
 Arg Arg Lys Leu Ser Arg Leu Pro Ser Tyr Ile Glu Ala Arg Arg Ala
 245 250 255
 Val Ala Arg Arg Tyr Glu Glu Gly Leu Ala Asp Thr Gly Leu Leu Leu
 260 265 270
 Pro Arg Thr Ala Gln Gly Asn Glu His Val Tyr Tyr Val Tyr Val Val
 275 280 285
 Arg His Pro Arg Arg Asp Ala Val Leu Glu Ala Leu Arg Ala Ser Tyr
 290 295 300
 Asp Ile Ala Leu Asn Ile Ser Tyr Pro Trp Pro Val His Thr Met Thr
 305 310 315 320
 Gly Phe Ser His Leu Gly Tyr Ala Lys Gly Ser Leu Pro Val Thr Glu
 325 330 335
 Ala Leu Ala Asp Glu Ile
 340

<210> 3
 <211> 293
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snogJ*, function: dTDP-glucose synthase"

<400> 3

Val Lys Gly Ile Ile Leu Ala Gly Gly Thr Gly Ser Arg Leu His Pro
 1 5 10 15
 Thr Thr Leu Ala Val Ser Lys Gln Leu Leu Pro Val Gly Asp Lys Pro
 20 25 30
 Met Ile Tyr Tyr Pro Leu Ser Val Leu Met Leu Ala Gly Val Thr Asp
 35 40 45
 Ile Leu Ile Ile Ser Thr Pro His Glu Leu Pro Arg Met Arg Arg Leu
 50 55 60
 Phe Gly Asp Gly Ala Gln Leu Gly Leu Arg Leu Ala Tyr Ala Glu Gln
 65 70 75 80
 Glu Lys Pro Arg Gly Ile Ala Glu Ala Phe Leu Ile Gly Ala Asp His
 85 90 95
 Val Gly Ser Asp Ala Val Ala Leu Ala Leu Gly Asp Asn Ile Phe His
 100 105 110
 Gly Ser Ser Phe Gln Gly Val Leu Arg Lys Glu Ala Glu Glu Leu Asp
 115 120 125
 Gly Cys Val Leu Phe Gly Tyr Pro Val Lys Asp Pro Gln Arg Tyr Gly
 130 135 140

11

Val Gly Glu Ala Asn Ala Ser Gly Arg Leu Val Ser Ile Glu Glu Lys
 145 150 155 160
 Pro Val Arg Pro Arg Ser Asn Arg Ala Ile Thr Gly Leu Tyr Phe Tyr
 165 170 175
 Asp Asn Glu Val Val Asp Ile Ala Arg Arg Leu Arg Pro Ser Ala Arg
 180 185 190
 Gly Glu Leu Glu Ile Thr Asp Ile Asn Arg Thr Tyr Met Glu Arg Gly
 195 200 205
 Arg Ala Arg Leu Val Asp Leu Gly Arg Gly Phe Ala Trp Leu Asp Thr
 210 215 220
 Gly Thr Pro Glu Ser Leu Leu Gln Ala Ser Gln Tyr Val Ser Ala Leu
 225 230 235 240
 Glu Glu Arg Gln Gly Ile Arg Ile Ala Cys Ile Glu Glu Val Ala Leu
 245 250 255
 Arg Met Gly Phe Ile Asn Ala Gln Ala Cys Tyr Glu Leu Gly Ala Arg
 260 265 270
 Leu Ser Gly Ser Gly Tyr Gly Gln Tyr Val Met Ala Ile Ala Glu Glu
 275 280 285
 Cys Thr Gly Arg Val
 290

<210> 4
 <211> 238
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snogA*, function: aminomethyl transferase"

<400> 4

Val Tyr Gly Arg Glu Leu Ala Asp Val Tyr Glu Met Val Tyr Arg Ser
 1 5 10 15
 Arg Gly Lys Ser Trp Ala Asp Glu Ala Glu Arg Val Thr Ala Glu Ile
 20 25 30
 Arg Ser Arg Arg Pro Gly Ala Arg Ser Leu Leu Asp Val Ala Cys Gly
 35 40 45
 Thr Gly Ala His Leu Glu Ala Phe Arg Gly Leu Phe Ala His Thr Glu
 50 55 60
 Gly Leu Glu Leu Ser Asp Glu Met Arg Ala Leu Ala Glu Arg Arg Leu
 65 70 75 80
 Pro Gly Val Pro Val Arg Pro Gly Asp Met Arg Asp Phe Ala Leu Ser
 85 90 95
 Gly Arg Phe Asp Ala Val Val Cys Leu Phe Cys Ser Ile Gly Tyr Leu
 100 105 110
 Glu Thr Val Ala Asp Met Arg Ala Ala Val Arg Thr Met Ala Ala His
 115 120 125
 Leu Val Pro Gly Gly Val Leu Val Val Glu Pro Trp Trp Phe Pro Glu
 130 135 140

12

Arg Phe Leu Glu Gly Tyr Val Ala Gly Asp Leu Ala Arg Gly Glu Gly
 145 150 155 160
 Arg Thr Val Ala Arg Val Ser His Ser Thr Arg Gln Gly Arg Arg Thr
 165 170 175
 Arg Met Glu Val Arg Phe Leu Val Gly Glu Ala Thr Gly Ile Arg Glu
 180 185 190
 Phe Thr Glu Ile Asp Leu Leu Thr Leu Phe Thr Arg Glu Glu Tyr Leu
 195 200 205
 Ala Ala Phe Glu Asp Ala Gly Cys Pro Ala Glu Phe Leu Asp Asp Gly
 210 215 220
 Leu Thr Gly Arg Gly Leu Phe Val Gly Val Arg Gly Ala Gly
 225 230 235

<210> 5
 <211> 324
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snoaM*, function: polyketide cyclase"

<400> 5

Met Thr Ala Ala Trp Gly Ala Pro Leu Tyr Pro Pro Trp Ile Pro Ala
 1 5 10 15
 Arg Pro Gly Arg Arg Arg Cys Gly Ala Gly Arg Arg Val Arg Cys Pro
 20 25 30
 Pro Val Glu Pro Ala Ser Arg Pro Arg Gln Glu Gly Arg Val Ser Val
 35 40 45
 Val Pro Ala Leu Arg Gln Pro Ser Pro Ser Thr Asn Pro Glu Val Arg
 50 55 60
 Val Arg Leu Ile Asp Leu Ser Ser Pro Val Asp Ser Ser Gln Tyr Glu
 65 70 75 80
 Pro Asp Pro Val Val His Asp Val Leu Thr Pro Arg Gln Gly Ala Glu
 85 90 95
 His Met Cys Ala Glu Met Arg Glu His Phe Gly Val Glu Phe Ser Pro
 100 105 110
 Asp Glu Leu Pro Asp Gly Glu Phe Leu Ser Leu Asp Arg Ile Thr Leu
 115 120 125
 Thr Thr His Thr Gly Thr His Val Asp Ala Pro Ser His Tyr Gly Ser
 130 135 140
 Arg Ala Leu Tyr Gly Asp Gly Val Pro Arg His Ile Asp Gln Met Pro
 145 150 155 160
 Leu Glu Trp Phe Phe Gly Arg Gly Val Val Leu Asp Leu Thr Asp Ala
 165 170 175
 Pro Thr Gly Thr Val Ser Ala Ala Arg Leu Glu Lys Glu Leu Ala Arg
 180 185 190
 Thr Gly Cys Ala Leu Arg Pro Gly Asp Ile Val Leu Leu His Thr Gly
 195 200 205

13

Ala Gln Arg His Ala Gly Thr Pro Arg Tyr Phe Thr Asp Phe Ala Gly
 210 215 220

Leu Asp Gly Pro Ala Val Arg Met Leu Leu Asp His Gly Val Arg Val
 225 230 235 240

Ile Gly Thr Asp Ala Phe Ser Leu Asp Ala Pro Phe Gly His Ile Ile
 245 250 255

Asp Arg Tyr Arg Ala Thr Gly Asp Arg Ser Val Leu Trp Pro Ala His
 260 265 270

Val Val Gly Arg Glu Arg Glu Tyr Cys Gln Ile Glu Arg Leu Ala Asn
 275 280 285

Leu Asp Arg Leu Pro Val Ser Phe Gly Phe Arg Val Cys Cys Phe Pro
 290 295 300

Val Lys Val Ala Gly Ala Gly Ala Gly Trp Thr Arg Ala Val Ala Leu
 305 310 315 320

Val Asp Glu Asp

<210> 6
 <211> 408
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snogN*, function: unknown"

<400> 6

Met Val Met Lys Leu Thr Asp Ser Glu Leu Gly Arg Ala Leu Leu Ser
 1 5 10 15

Leu Arg Gly Tyr Gln Trp Leu Arg Gly Ile His His Asp Pro Tyr Ala
 20 25 30

Leu Leu Leu Arg Ala Glu Ser Asp Asp Pro Ala Gln Leu Gly Arg Leu
 35 40 45

Leu Arg Glu Arg Gly Arg Leu His Arg Ser Asp Thr Gly Thr Trp Val
 50 55 60

Thr Ala Asp His Ala Thr Ala Ser Arg Leu Leu Ala Asp Pro Arg Phe
 65 70 75 80

Val Leu Arg Arg Pro Ala Gly Pro Ala Thr Gly Thr Gly Asp Val
 85 90 95

Met Pro Trp Glu Glu Ala Thr Leu Ser Asp Leu Leu Pro Leu Asp Glu
 100 105 110

Ala Arg Leu Thr Thr Asp Arg Ala Arg Cys Arg Arg Leu Gly Ala Thr
 115 120 125

Ala Ala Arg Ile Ala Ala Asp Gly Pro Val Ala Thr Arg Leu Ala Asp
 130 135 140

Leu Ala Gly Ala Arg Ala Glu Gln Val Arg Ser Thr Gly His Phe Asp
 145 150 155 160

Leu Arg Ala Asp Tyr Ala Leu Pro Tyr Ala Val Glu Pro Ala Cys Ala
 165 170 175

14

Leu Leu Gly Leu Pro Ala Gly Gln Cys Ser Leu Phe Gly Ala Phe Ser
 180 185 190
 Pro Ala Val Leu Leu Asp Ala Thr Val Val Pro Pro Arg Leu Pro Glu
 195 200 205
 Ala Arg Ala Leu Ile Ala Ser Thr Ala Glu Leu Thr Ala Leu Trp Pro
 210 215 220
 Arg Leu Ala Pro Ser Leu Ser Lys Thr Val Pro Glu Asp Glu Ala Pro
 225 230 235 240
 Asp Leu Phe Leu Leu Thr Ala Val Leu Leu Val Pro Ala Val Val His
 245 250 255
 Leu Val Cys Glu Ala Val Ala Ala Leu Ser His Asp Pro Gly Gln Ala
 260 265 270
 Gly Leu Leu Arg Asp Asp Pro Val Leu Ala Ala Pro Ala Val Glu Glu
 275 280 285
 Thr Leu Arg His Ala Pro Pro Ala Arg Leu Phe Thr Leu His Ala Thr
 290 295 300
 Gly Pro Glu Arg Val Ala Asp Val Asp Leu Pro Ala Gly Ala Glu Val
 305 310 315 320
 Ala Val Val Val Ala Ala Ala His Arg Asp Pro Ser Trp Cys Pro Asp
 325 330 335
 Pro Asp Arg Phe Asp Leu Thr Arg Asn Glu Arg His Leu Ala Leu Pro
 340 345 350
 Pro Asp Leu Pro Leu Gly Ala Leu Ala Pro Leu Leu Arg Val Cys Ala
 355 360 365
 Thr Ala Ala Val Ala Ala Leu Ala Ala Gly Leu Leu Pro Leu Arg Ala
 370 375 380
 Val Gly Pro Pro Val Arg Arg Leu Arg Ala Pro Val Thr Arg Ser Val
 385 390 395 400
 Leu Arg Phe Pro Val Ala Pro Cys
 405

<210> 7
 <211> 422
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451
 <220>
 <223> "translate of *snoaG*, function: hydroxylase"
 <400> 7

Met Asp Asn Arg Glu Thr Val Arg Pro Val Ser Val Cys Arg Val Cys
 1 5 10 15
 Gly Gly Asn Asp Trp Gln Asp Val Val Asp Phe Gly Asp Val Pro Leu
 20 25 30
 Ala Asn Gly Phe Leu Ser Pro Ala Asp Ser Tyr Glu Asn Glu Arg Arg
 35 40 45
 Tyr Pro Leu Gly Val Leu Ser Cys Arg Ala Cys Arg Leu Met Ser Leu
 50 55 60

15

Thr His Val Val Asp Pro Glu Val Leu Tyr Arg Asp Tyr Ala Tyr Thr
 65 70 75 80
 Thr Pro Asp Ser Glu Met Ile Thr Gln His Met Arg His Ile Thr Ala
 85 90 95
 Leu Cys Arg Thr Arg Phe Glu Leu Pro Pro Asp Ser Leu Val Val Glu
 100 105 110
 Leu Gly Ser Asn Thr Gly Arg Gln Leu Met Ala Phe Arg Glu Ala Gly
 115 120 125
 Met Arg Thr Leu Gly Val Asp Pro Ala Arg Asn Leu Thr Asp Val Ala
 130 135 140
 Arg Arg Asn Gly Ile Glu Thr Phe Pro Asp Phe Phe Ser His Asp Val
 145 150 155 160
 Ala Arg Thr Ile Arg Arg Asp His Gly Gln Ala Arg Leu Val Leu Gly
 165 170 175
 Arg His Val Phe Ala His Ile Asp Asp Val Ser Asp Ile Ala Ala Gly
 180 185 190
 Val Arg Glu Leu Leu Ser Pro Asp Gly Val Phe Ala Ile Glu Val Pro
 195 200 205
 Tyr Val Leu Asp Leu Leu Glu Lys Val Ala Phe Asp Thr Ile Tyr His
 210 215 220
 Glu His Leu Ser Tyr Phe Thr Met Arg Ser Phe Val Thr Leu Phe Ala
 225 230 235 240
 Arg His Gly Leu Arg Val Leu Asp Val Glu Arg Phe Gly Val His Gly
 245 250 255
 Gly Ser Val Leu Val Phe Val Gly His Glu Asp Gly Pro Trp Pro Glu
 260 265 270
 Arg Pro Ser Val Pro Glu Leu Leu Arg Val Glu Arg Gln Arg Gly Leu
 275 280 285
 Tyr Asp Asp Ala Thr Tyr Arg Thr Phe Ala Gln Arg Ile Glu Arg Val
 290 295 300
 Arg Thr Glu Leu Pro Glu Leu Leu Arg Ser Leu Val Ala Gln Gly Lys
 305 310 315 320
 Arg Ile Val Gly Tyr Gly Ala Pro Ala Lys Gly Asn Thr Ile Leu Thr
 325 330 335
 Val Cys Gly Leu Gly Leu Lys Glu Leu Glu Tyr Cys Thr Asp Thr Thr
 340 345 350
 Glu Leu Lys Gln Gly Arg Val Leu Pro Gly Thr His Ile Pro Val His
 355 360 365
 Ala Pro Glu His Ala Lys Glu His Ile Pro Asp Tyr Tyr Leu Leu Leu
 370 375 380
 Ala Trp Asn Tyr Ala Thr Glu Ile Leu Asp Lys Glu Thr Ala Phe Arg
 385 390 395 400
 Asp Asn Gly Gly Arg Phe Ile Val Pro Ile Pro Arg Pro Ser Ile Leu
 405 410 415

Thr Ser Pro Ser Gly Ser
420

<210> 8
 <211> 291
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snogC*, function: dTDP-4-dehydrorhamnose reductase"

<400> 8

Met Leu Ala Arg His Leu Thr Ala Ala Leu Ala Glu Thr Gly Arg Ser
1 5 10 15

Arg Pro Ala Ala Glu Ala Val Val Leu Gly Arg Arg Ala Leu Asp Ile
20 25 30

Thr Asp Gly Arg Ala Val Asp Ala Ala Phe Ala Ala His Arg Pro Arg
35 40 45

Val Val Val Asn Cys Ala Ala Phe Thr Asp Val Asp Gly Ala Glu Ser
50 55 60

Arg Trp Ala Glu Ala Met Arg Val Asn Gly Gly Gly Pro Arg Leu Leu
65 70 75 80

Ala Arg Arg Cys Ala Arg His Gly Val Arg Leu Ile His Val Ser Thr
85 90 95

Asp Tyr Val Phe Pro Gly Asp Thr Arg Ser Pro Tyr Gly Glu Ser Asp
100 105 110

Ala Pro Gly Pro Arg Thr Val Tyr Gly Arg Ser Lys Leu Ala Gly Glu
115 120 125

Arg Ala Val Leu Ser Leu Leu Pro Asp Thr Gly Thr Val Val Arg Thr
130 135 140

Ala Trp Leu Tyr Gly Gly Gln Gly Arg Ser Phe Val Arg Thr Met Leu
145 150 155 160

Glu Arg Ala Pro Asp Asp Gly His Val Asp Val Val Asn Asp Gln Trp
165 170 175

Gly Gln Pro Thr Trp Ala Gly Asp Val Ala Arg Leu Leu Val Thr Leu
180 185 190

Ala Arg Thr Pro Pro Asp Arg Ala Arg Gly Ile Phe His Ala Thr Asn
195 200 205

Ala Gly Ala Ala Thr Trp Tyr Glu Leu Ala Arg Glu Val Phe Arg Leu
210 215 220

Ala Gly Ala Asp Pro Glu Arg Val Arg Pro Val Ala Thr Ala Asp Arg
225 230 235 240

Pro Gly Pro Ala Pro Arg Pro Ala Cys Thr Val Leu Gly His Asp Arg
245 250 255

Trp Arg Leu Val Gly Val Ala Pro Pro Arg Asp Trp Arg Ala Ala Leu
260 265 270

Arg Glu Ala Met Arg Gln Leu Leu Pro Gly Gly Arg Leu Arg Asn Leu
275 280 285

Thr Gly Thr
 290

<210> 9
 <211> 350
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snogK*, function: dTDP-glucose-4,6-dehydratase"

<400> 9

Met Ala Ser His Thr Ser Ala Thr Thr Asp Val Asn Ile Leu Val Thr
 1 5 10 15
 Gly Ala Val Gly Phe Ile Gly Ser Ala Tyr Val Arg Met Leu Leu Glu
 20 25 30
 Asn Arg Ala Pro Gly Ala Gly Ala Pro Ala Val Arg Val Thr Val Leu
 35 40 45
 Asp Lys Leu Thr Tyr Ala Gly Asn Leu Thr Asn Leu Asp Ala Val Arg
 50 55 60
 Gly Asp Arg Leu Arg Phe Val Arg Gly Asp Ile Leu Asp Ala Glu Leu
 65 70 75 80
 Val Asp Glu Leu Met Ala His Ser Asp Gln Val Val His Phe Ala Ala
 85 90 95
 Glu Ser His Val Asp Arg Ser Ile Arg Ala Ala Asp Asp Phe Val Leu
 100 105 110
 Thr Asn Val Val Gly Thr Gln Arg Leu Leu Asp Ala Ala Leu Arg His
 115 120 125
 Gly Val Glu Pro Phe Val Leu Val Ser Thr Asp Glu Val Tyr Gly Ser
 130 135 140
 Ile Ala Ser Gly Ser Trp Pro Glu Glu His Pro Leu Ser Pro Asn Ser
 145 150 155 160
 Pro Tyr Ala Ala Ser Lys Ala Ser Ala Asp Leu Met Ala Phe Ala Cys
 165 170 175
 His Arg Thr His Gly Leu Asp Val Arg Val Thr Arg Cys Ser Asn Asn
 180 185 190
 Tyr Gly Pro Arg Gln His Pro Glu Lys Leu Ile Pro Arg Phe Val Thr
 195 200 205
 Asn Leu Leu Asp Gly Leu Pro Val Pro Leu Tyr Gly Asp Gly Arg Asn
 210 215 220
 Val Arg Glu Trp Leu His Val Glu Asp His Cys Arg Gly Val Asp Leu
 225 230 235 240
 Val Arg Thr Ala Gly Arg Pro Gly Gly Val Tyr His Ile Gly Gly Gly
 245 250 255
 Arg Glu Leu Ser Asn Arg Glu Leu Val Gly Met Leu Leu Glu Leu Cys
 260 265 270
 Gly Ala Asp Trp Ser Ser Val Arg His Val Pro Asp Arg Lys Gly His
 275 280 285

18

Asp Leu Arg Tyr Ser Leu Asp Trp Gly Arg Ala Arg Glu Glu Leu Gly
 290 295 300
 Tyr Arg Pro Ala Arg Glu Phe Ser Ser Gly Leu Arg Ser Thr Val Gln
 305 310 315 320
 Trp Tyr Arg Glu Asn Arg Ser Trp Trp Glu Pro Leu Lys Arg Gly Val
 325 330 335
 Thr Ala Pro Gly Gly Thr Ser Thr Val Val Pro Gly Val Arg
 340 345 350

<210> 10
 <211> 134
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snoaL*, function: NAME cyclase"

<400> 10

Met Val Ser Ala Phe Asn Thr Gly Arg Thr Asp Asp Val Asp Glu Tyr
 1 5 10 15
 Ile His Pro Asp Tyr Leu Asn Pro Ala Thr Leu Glu His Gly Ile His
 20 25 30
 Thr Gly Pro Lys Ala Phe Ala Gln Leu Val Gly Trp Val Arg Ala Thr
 35 40 45
 Phe Ser Glu Glu Ala Arg Leu Glu Glu Val Arg Ile Glu Glu Arg Gly
 50 55 60
 Pro Trp Val Lys Ala Tyr Leu Val Leu Tyr Gly Arg His Val Gly Arg
 65 70 75 80
 Leu Val Gly Met Pro Pro Thr Asp Arg Arg Phe Ser Gly Glu Gln Val
 85 90 95
 His Leu Met Arg Ile Val Asp Gly Lys Ile Arg Asp His Arg Asp Trp
 100 105 110
 Pro Asp Phe Gln Gly Thr Leu Arg Gln Leu Gly Asp Pro Trp Pro Asp
 115 120 125
 Asp Glu Gly Trp Arg Pro
 130

<210> 11
 <211> 235
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snoK*, function: unknown"

<400> 11

Met Pro Asp Pro Gly Gly Pro Thr Thr Ala Glu Asn Leu Ser Lys Glu
 1 5 10 15
 Ala Val Arg Phe Tyr Arg Glu Gln Gly Tyr Val His Ile Pro Arg Val
 20 25 30
 Leu Ser Glu Thr Glu Val Thr Ala Phe Arg Ala Ala Cys Glu Glu Val
 35 40 45

19

Leu Glu Lys Glu Gly Arg Glu Ile Ser Gly Ile Ala Leu Arg Leu Ala
 50 55 60
 Gly Ala Pro Leu Arg Val Tyr Ser Ser Asp Ile Leu Val Lys Glu Pro
 65 70 75 80
 Lys Arg Thr Leu Pro Thr Leu Val His Asp Asp Glu Thr Gly Leu Pro
 85 90 95
 Leu Asn Glu Leu Ser Ala Thr Leu Thr Ala Trp Ile Ala Leu Thr Asp
 100 105 110
 Val Pro Val Glu Arg Gly Cys Met Ser Tyr Val Pro Gly Ser His Leu
 115 120 125
 Arg Ala Arg Glu Asp Arg Gln Glu His Met Thr Ser Phe Ala Glu Phe
 130 135 140
 Arg Asp Leu Ala Asp Val Trp Pro Asp Tyr Pro Trp Gln Pro Arg Val
 145 150 155 160
 Ala Val Pro Val Arg Ala Gly Asp Val Val Phe His His Cys Arg Thr
 165 170 175
 Val His Met Ala Glu Ala Asn Thr Ser Asp Ser Val Arg Met Ala His
 180 185 190
 Gly Val Val Tyr Met Asp Ala Asp Ala Thr Tyr Arg Pro Gly Val Gln
 195 200 205
 Asp Gly His Leu Ser Arg Leu Ser Pro Gly Asp Pro Leu Glu Gly Glu
 210 215 220
 Leu Phe Pro Leu Val Thr Ala Gly Thr Arg Gln
 225 230 235

<210> 12

<211> 390

<212> PRT

<213> *Streptomyces nogalater* ATCC 27451

<220>

<223> "translate of *snogD*, function: glycosyl transferase"

<400> 12

Met Arg Val Pro Gly Ser Cys Arg Thr Gly Gly Ile Met Arg Ala Leu
 1 5 10 15
 Phe Ile Thr Ser Pro Gly Leu Ser His Ile Leu Pro Thr Val Pro Leu
 20 25 30
 Ala Gln Ala Leu Arg Ala Leu Gly His Glu Val Arg Tyr Ala Thr Gly
 35 40 45
 Gly Asp Ile Arg Ala Val Ala Glu Ala Gly Leu Cys Ala Val Asp Val
 50 55 60
 Ser Pro Gly Val Asn Tyr Ala Lys Leu Phe Val Pro Asp Asp Thr Asp
 65 70 75 80
 Val Thr Asp Pro Met His Ser Glu Gly Leu Gly Glu Gly Phe Phe Ala
 85 90 95
 Glu Met Phe Ala Arg Val Ser Ala Val Ala Val Asp Gly Ala Leu Arg
 100 105 110

20

Thr Ala Arg Ser Trp Arg Pro Asp Leu Val Val His Thr Pro Thr Gln
 115 120 125
 Gly Ala Gly Pro Leu Thr Ala Ala Ala Leu Gln Leu Pro Cys Val Glu
 130 135 140
 Leu Pro Leu Gly Pro Ala Asp Ser Glu Pro Gly Leu Gly Ala Leu Ile
 145 150 155 160
 Arg Arg Ala Met Ser Lys Asp Tyr Glu Arg His Gly Val Thr Gly Glu
 165 170 175
 Pro Thr Gly Ser Val Arg Leu Thr Thr Thr Pro Pro Ser Val Glu Ala
 180 185 190
 Leu Leu Pro Glu Asp Arg Arg Ser Pro Gly Ala Trp Pro Met Arg Tyr
 195 200 205
 Val Pro Tyr Asn Gly Gly Ala Val Leu Pro Asp Trp Leu Pro Pro Ala
 210 215 220
 Ala Gly Arg Arg Arg Ile Ala Val Thr Leu Gly Ser Ile Asp Ala Leu
 225 230 235 240
 Ser Gly Gly Ile Ala Lys Leu Ala Pro Leu Phe Ser Glu Val Ala Asp
 245 250 255
 Val Asp Ala Glu Phe Val Leu Thr Leu Gly Gly Gly Asp Leu Ala Leu
 260 265 270
 Leu Gly Glu Leu Pro Ala Asn Val Pro Val Val Glu Trp Ile Pro Leu
 275 280 285
 Gly Ala Leu Leu Glu Thr Cys Asp Ala Ile Ile His His Gly Gly Ser
 290 295 300
 Gly Thr Leu Leu Thr Ala Leu Ala Ala Gly Val Pro Gln Cys Val Ile
 305 310 315 320
 Pro His Gly Ser Tyr Gln Asp Thr Asn Arg Asp Val Leu Thr Gly Leu
 325 330 335
 Gly Ile Gly Phe Asp Ala Glu Ala Gly Ser Leu Gly Ala Glu Gln Cys
 340 345 350
 Arg Arg Leu Leu Asp Asp Ala Gly Leu Arg Glu Ala Ala Leu Arg Val
 355 360 365
 Arg Gln Glu Met Ser Glu Met Pro Pro Pro Ala Glu Thr Ala Ala Lys
 370 375 380
 Leu Val Ala Leu Ala Gly
 385 390

<210> 13
 <211> 275
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snoW*, function: unknown"

<400> 13

Met Thr Val Leu Val Thr Gly Ala Thr Gly Asn Val Gly Arg His Val
 1 5 10 15

21

Val Thr Gly Leu Leu Ala Ala Gly Arg Arg Val Arg Ala Leu Thr Arg
 20 25 30
 Thr Pro Asp Arg Ser Gly Leu Pro Gly Gly Ala Glu Ile Thr Gly Gly
 35 40 45
 Asp Leu Thr Arg Pro Glu Thr Tyr Glu Arg Met Leu Asp Gly Val Glu
 50 55 60
 Ala Val Tyr Leu Phe Pro Val Pro Glu Thr Ala Ala Ala Phe Ala Gly
 65 70 75 80
 Ala Ala Arg Arg Ala Gly Val Arg Arg Ile Val Val Leu Ser Ser Asp
 85 90 95
 Ser Val Thr Asp Gly Thr Asp Thr Gly Gly His Arg Arg Val Glu Leu
 100 105 110
 Ala Val Glu Asp Thr Gly Leu Glu Trp Thr His Val Arg Pro Gly Glu
 115 120 125
 Phe Ala Leu Asn Lys Val Thr Leu Trp Ala Pro Ser Ile Arg Ala Glu
 130 135 140
 Gly Val Val Arg Ser Ala Tyr Pro Asp Ala Arg Val Ala Pro Val His
 145 150 155 160
 Glu Ala Asp Val Ala Ala Val Ala Val Thr Ala Leu Leu Lys Glu Gly
 165 170 175
 His Ala Gly Arg Ala Tyr Ser Val Thr Gly Pro Gln Ala Leu Thr Gln
 180 185 190
 Arg Glu Gln Val Arg Ala Val Gly Glu Gly Leu Gly Arg Ser Leu Ala
 195 200 205
 Phe Val Glu Val Thr Pro Gly Gln Ala Arg Ala Asp Leu Thr Ala Gln
 210 215 220
 Gly Leu Pro Ala Pro Ile Ala Asp Tyr Val Leu Ala Phe Gln Ala Gly
 225 230 235 240
 Trp Thr Glu Arg Pro Ala Pro Ala Arg Pro Thr Val Arg Glu Val Thr
 245 250 255
 Gly Arg Pro Ala Arg Thr Leu Ala Gln Trp Ala Ala Asp His Arg Ala
 260 265 270
 Asp Phe Arg
 275

<210> 14
 <211> 424
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snogE*, function: glycosyl transferase"
 <400> 14

Val Arg Val Leu Leu Thr Ser Phe Ala Met Asp Ala His Phe Cys Thr
 1 5 10 15
 Ala Val Pro Leu Ala Trp Ala Leu Arg Ser Ala Gly His Glu Val Arg
 20 25 30

Val Ala Gly Gln Pro Ala Leu Thr Ser Thr Ile Thr Gly Ala Gly Leu
 35 40 45
 Thr Ala Val Pro Val Gly Arg Asp His Thr His Gly Ser Leu Leu Gly
 50 55 60
 Arg Val Gly Ser Asp Ile Leu Ala Leu His Asp Glu Ala Asp Tyr Leu
 65 70 75 80
 Glu Ala Arg His Asp Ala Leu Gly Phe Glu Phe Leu Lys Gly His Asn
 85 90 95
 Thr Val Met Ser Ala Leu Phe Tyr Ser Gln Ile Asn Asn Asp Ser Met
 100 105 110
 Val Asp Asp Leu Val Asp Phe Ala Arg His Trp Arg Pro Asp Leu Val
 115 120 125
 Val Trp Glu Pro Phe Thr Phe Ala Gly Ala Val Ala Ala Arg Ala Ser
 130 135 140
 Gly Ala Ala His Ala Arg Leu Leu Ser Phe Pro Asp Leu Phe Leu Ser
 145 150 155 160
 Thr Arg Arg Leu Phe Leu Glu Arg Met Ala Arg Gln Glu Pro Glu His
 165 170 175
 His Asp Asp Thr Leu Ala Glu Trp Leu Asp Trp Thr Leu Gly Arg His
 180 185 190
 Gly His Ser Phe Asp Glu Glu Ile Val Thr Gly Gln Trp Ser Ile Asp
 195 200 205
 Gln Thr Pro Ala Pro Val Arg Leu Asp Ala Gly Gly Pro Thr Val Pro
 210 215 220
 Met Arg Tyr Val Pro Tyr Ser Gly Leu Val Pro Thr Val Val Pro Asp
 225 230 235 240
 Trp Leu Arg Arg Pro Pro Glu Arg Pro Arg Val Leu Val Thr Leu Gly
 245 250 255
 Ile Thr Ser Arg Arg Val Lys Ser Phe Leu Ala Val Ser Val Asp Asp
 260 265 270
 Leu Phe Glu Ala Val Ala Gly Leu Gly Val Glu Val Val Ala Thr Leu
 275 280 285
 Asp Ala Asp Gln Arg Glu Leu Leu Gly Arg Val Pro Asp His Phe Arg
 290 295 300
 Ile Val Glu His Val Pro Leu Asp Ala Val Leu Pro Thr Cys Ser Ala
 305 310 315 320
 Ile Val His His Gly Gly Ala Gly Thr Trp Ser Thr Ala Ala Val Tyr
 325 330 335
 Gly Val Pro Gln Val Ser Leu Gly Ser Met Trp Asp His Phe Tyr Arg
 340 345 350
 Ala Arg Arg Leu Glu Glu Leu Gly Ala Gly Leu Arg Leu Pro Ser Gly
 355 360 365
 Glu Leu Thr Ala Glu Gly Leu Arg Thr Arg Leu Glu Arg Val Leu Gly
 370 375 380

23

Glu Pro Ser Phe Gly Thr Ala Ala Gln Ala Leu Ser Asp Thr Ile Ala
385 390 395 400

Ala Glu Pro Ser Pro Ser Glu Val Val Pro Val Leu Glu Glu Leu Thr
405 410 415

Gly Arg His Arg Pro Gly Thr Arg
420

<210> 15
<211> 139
<212> PRT
<213> *Streptomyces nogalater* ATCC 27451

<220>
<223> "translate of *snoL*, function: unknown"

<400> 15
Met Ser Thr Thr Ala Asn Lys Glu Arg Cys Leu Glu Met Val Ala Ala
1 5 10 15
Trp Asn Arg Trp Asp Val Ser Gly Val Val Ala His Trp Ala Pro Asp
20 25 30
Val Val His Tyr Asp Asp Glu Asp Lys Pro Val Ser Ala Glu Glu Val
35 40 45
Val Arg Arg Met Asn Ser Ala Val Glu Ala Phe Pro Asp Leu Arg Leu
50 55 60
Asp Val Arg Ser Ile Val Gly Glu Gly Asp Arg Val Met Leu Arg Ile
65 70 75 80
Thr Cys Ser Ala Thr His Gln Gly Val Phe Met Gly Ile Ala Pro Thr
85 90 95
Gly Arg Lys Val Arg Trp Thr Tyr Leu Glu Glu Leu Arg Phe Ser Glu
100 105 110
Ala Gly Lys Val Val Glu His Trp Asp Val Phe Asn Phe Ser Pro Leu
115 120 125
Phe Arg Asp Leu Gly Val Val Pro Asp Gly Leu
130 135

<210> 16
<211> 155
<212> PRT
<213> *Streptomyces nogalater* ATCC 27451

<220>
<223> "translate of *snoO*, function: homologous to *mtmX* of mithramycin cluster"

<400> 16
Met Ser Val Arg Thr Asp Gln Thr Ala Ala Pro Glu Asp Arg Ala Ala
1 5 10 15
Ala Thr Asp Pro Gly Phe Gly His Leu Tyr Ala Gln Val Gln Gln Phe
20 25 30
Tyr Ala Arg Gln Met Gln Leu Leu Asp Ser Gly Ala Ala Glu Glu Trp
35 40 45

24

Ala Ala Thr Phe Thr Glu Asp Gly Thr Phe Ala Arg Pro Ser Ser Pro
 50 55 60
 Glu Pro Ala Arg Gly His Ala Glu Leu Ala Ala Gly Ala Arg Ala Ala
 65 70 75 80
 Ala Glu Arg Leu Ala Ala Glu Gly Leu Ser His Arg His Val Ile Gly
 85 90 95
 Met Thr Ala Val Arg Arg Glu Pro Asp Gly Ser Val Phe Val Arg Ser
 100 105 110
 Tyr Ala Gln Val Phe Ala Thr Arg Arg Gly Glu Ala Pro Arg Leu His
 115 120 125
 Leu Ile Cys Val Cys Glu Asp Val Leu Val Arg Glu Gly Pro Gly Leu
 130 135 140
 Lys Val Arg Glu Arg Val Val Thr His Asp Ala
 145 150 155

<210> 17
 <211> 281
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snoaF*, function: C-7 ketoreductase"

<400> 17

Val Arg Ala Met Thr Asp Ser Thr Gly Pro Arg Pro Val Pro Ala Met
 1 5 10 15
 Ser Pro Ala Pro Ser Pro Thr Pro Ser Pro Gly Pro Ala Pro Gly Ser
 20 25 30
 Glu Pro Ala Pro Leu Ala Val Ile Val Thr Gly Gly Gly Ser Gly Ile
 35 40 45
 Gly Arg Ala Thr Ala Arg Ala Phe Ala Ala Gln Gly Ala Lys Val Leu
 50 55 60
 Val Val Gly Arg Thr Glu Asp Ala Leu Ala Gln Thr Ala Glu Gly Cys
 65 70 75 80
 Ala Asp Met Arg Val Leu Val Ala Asp Val Ala Ser Pro Asp Gly Pro
 85 90 95
 Gln Ala Val Val Asn Ala Ala Leu Arg Glu Phe Gly Arg Ile Asp Val
 100 105 110
 Leu Val Asn Asn Ala Ala Val Ala Gly Met Glu Thr Leu Gln Thr Val
 115 120 125
 Asp Arg Asp Ala Val Ala Arg Gln Phe Gly Thr Asn Leu Thr Ala Pro
 130 135 140
 Leu Phe Leu Val Gln Ser Ala Leu Gly Ala Leu Glu Lys Ser Arg Gly
 145 150 155 160
 Ile Val Val Asn Val Gly Thr Ala Ala Thr Leu Gly Leu Arg Ala Ala
 165 170 175
 Pro Thr Gly Ala Leu Tyr Gly Ala Ser Lys Val Ala Leu Asp Tyr Leu
 180 185 190

25

Thr Arg Thr Trp Ala Val Glu Leu Ala Pro Arg Gly Ile Arg Val Val
 195 200 205

Gly Val Ala Pro Gly Val Ile Asp Thr Gly Ile Gly Val Arg Met Gly
 210 215 220

Met Thr Pro Glu Gly Tyr Arg Glu Phe Leu Thr Gly Met Gly Gly Arg
 225 230 235 240

Val Pro Val Gly Arg Val Gly Arg Pro Glu Asp Val Ala Trp Trp Ile
 245 250 255

Val Gln Leu Ala Arg Pro Glu Ala Gly Tyr Ala Thr Gly Met Val Val
 260 265 270

Pro Val Asp Gly Gly Leu Ser Leu Val
 275 280

<210> 18
 <211> 190
 <212> PRT
 <213> *Streptomyces nogalater* ATCC 27451

<220>
 <223> "translate of *snoN*, function: unknown"

<400> 18

Val Gln Glu Thr Glu Pro Gly Val Pro Ala Asp Leu Pro Ala Glu Ser
 1 5 10 15

Asp Pro Ala Ala Leu Glu Arg Leu Ala Ala Arg Tyr Arg Arg Asp Gly
 20 25 30

Tyr Val His Val Pro Gly Val Leu Asp Ala Gly Glu Val Ala Glu Tyr
 35 40 45

Leu Ala Glu Ala Arg Arg Leu Leu Ala His Glu Glu Ser Val Arg Trp
 50 55 60

Gly Ser Gly Ala Gly Thr Val Met Asp Tyr Val Ala Asp Ala Gln Leu
 65 70 75 80

Gly Ser Asp Thr Met Arg Arg Leu Ala Thr His Pro Arg Ile Ala Ala
 85 90 95

Leu Ala Glu Tyr Leu Ala Gly Ser Pro Leu Arg Leu Phe Lys Leu Glu
 100 105 110

Val Leu Leu Lys Glu Asn Lys Glu Lys Asp Ala Ser Val Pro Thr Ala
 115 120 125

Pro His His Asp Ala Phe Ala Phe Pro Phe Ser Thr Ala Gly Thr Ala
 130 135 140

Leu Thr Ala Trp Val Ala Leu Val Asp Val Pro Val Glu Arg Gly Cys
 145 150 155 160

Met Thr Phe Val Pro Gly Ser His Leu Leu Pro Asp Pro Asp Thr Gly
 165 170 175

Asp Glu Pro Trp Ala Gly Ala Phe Thr Arg Pro Gly Glu Ile
 180 185 190